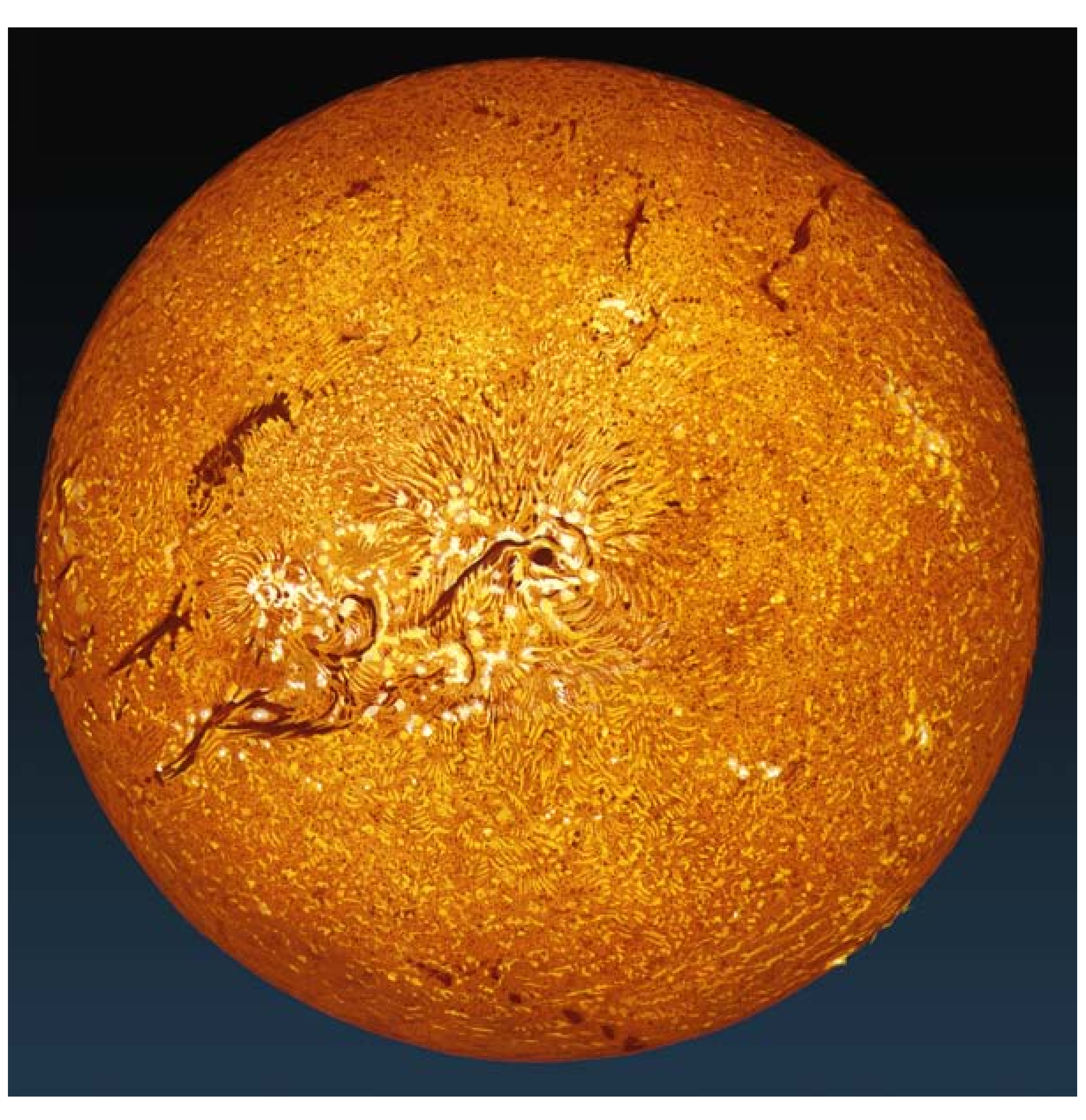
# INVESTIGACION y CIENCIA

CAOS Y FRACTALES EN LA FISIOLOGIA HUMANA

TRANSICIONES DE FASE EN LAS PEROVSKITAS

BIOQUIMICA DE LA FOTOSINTESIS

Edición española de SCIENTIFIC AMERICAN





CAMBIOS EN LA RADIACION SOLOAR

ABRIL 1990
500 PTAS.




Abril de 1990 Número 163

Edición española de **SCIENTIFIC AMERICAN**



## SECCIONES

## COLABORADORES DE ESTE NUMERO

**Asesoramiento y traducción:**

José Manuel García de la Mora: *La tragedia del dolor innecesario*; Manuel Puigcerver: *Sol cambiante*; Luis Bou: *Caos y fractales en la fisiología humana* y *Juegos matemáticas*; Manuel Losada, Miguel Angel de la Rosa y Manuel Hervás: *Cómo producen oxígeno las plantas*; Amando García: *Avances en semiconductores de arseniuro de galio*; Joan-Domènec Ros: *Así comparten su alimento los vampiros*; Francisco Bustelo: *Retroacción positiva en la economía*; J. Vilardell: *Taller y laboratorio*.

**Ciencia y sociedad:**

A. Sánchez Lavega, J. D. Ros y J. E. Llebot

**Ciencia y empresa:**

Manuel Puigcerver

**Libros:**

León Garzón, Pedro Molera, Luis Alonso y José L. Sánchez

## PROCEDENCIA DE LAS ILUSTRACIONES

| Página | Fuente |
|---|---|
| 12-13 | Patricia J. Wynne |
| 15 | Carol Donner |
| 17 | Linda Bartlett |
| 21 | Observatorio Solar Nacional, Nuevo México |
| 22 | Ian Worpole |
| 23 | Laboratorio de Investigación Naval, Washington, D.C. (*arriba*), Observatorio Solar Nacional, Kitt Peak, Arizona (*abajo*) |
| 24-26 | Ian Worpole |
| 27 | Peter V. Foukal (*izquierda*), NASA (*derecha*) |
| 30 | Quesada/Burke, cortesía de John Lehr |
| 31 | Yoichiro Kawaguchi |
| 32-33 | Carol Donner, elaborado a partir de *A Textbook of Histology*, de Don W. Fawcett, W. B. Saunders Co.,© |
| 34 | Burton Sobel, Escuela Universitaria de Medicina de Washington |
| 35 | Carol Donner |
| 36 | Edward Bell y Joseph Mietus |
| 37 | Carla J. Shatz, Escuela Universitaria de Medicina de Stanford |
| 51 | Quesada/Burke |
| 52-54 | George Retseck |
| 55 | Johnny Johnson |
| 56 | George Retseck |
| 59 | Randall M. Feenstra y Joseph A. Stroncio, IBM |
| 60-61 | George V. Kelvin |
| 62 | IBM |
| 64 | George V. Kelvin |
| 66 | Edward Bell |
| 69 | Gunter Ziesler, Peter Arnold, Inc. |
| 70 | M. W. Larson, Bruce Coleman Inc. (*arriba*), Patricia J. Wynne (*abajo*) |
| 71 | Gerald S. Wilkinson |
| 72-73 | Patricia J. Wynne |
| 74 | Merlin D. Tuttle |
| 77-85 | Juan Bartolomé, Rafael Navarro y Magda Màrìa |
| 87 | Andrew Christie |
| 88 | Casa Editrice Giusti di Becocci |
| 89-90 | Andres Christie |
| 91 | Paul Wallich |
| 94 | Jearl Walker |
| 95-98 | Michael Goodman |
| 99-106 | M. Izoard |



EL DIBUJO DE LA PORTADA, basado en una fotografía tomada en el Observatorio Solar Nacional de Sacramento, muestra la violenta agitación típica de la baja atmósfera del Sol durante las épocas de gran actividad. El número de brillantes fulguraciones y de filamentos oscuros crece y disminuye en un ciclo de 11 años, relacionado con variaciones del campo magnético solar [véase "Sol cambiante", por Peter V. Foukal, en este mismo número]. El ciclo altera también la emisión solar de radiación y partículas cargadas, que pueden influir en la climatología terrestre.

**INVESTIGACION Y CIENCIA**

DIRECTOR GENERAL Francisco Gracia Guillén
DIRECTOR EDITORIAL José María Valderas Gallardo
DIRECTORA DE ADMINISTRACIÓN Pilar Bronchal Garfella
PRODUCCIÓN César Redondo Zayas
SECRETARÍA Purificación Mayoral Martínez
EDITA Prensa Científica, S. A. Viladomat, 291 6º 1ª - 08029 Barcelona (ESPAÑA)
Teléfonos 321 81 91 - 321 83 48 Telefax 419 47 82

**SCIENTIFIC AMERICAN**

EDITOR Jonathan Piel
BOARD OF EDITORS Armand Schwab, Jr., *Managing Editor*; Timothy Appenzeller, Laurie Burnham, *Associate Editors*; Timothy M. Beardsley; Elizabeth Corcoran; John Horgan; June Kinoshita; Philip Morrison, *Book Editor*; John Rennie; Philip E. Ros; Ricki L. Rusting; Russell Ruthen, Paul Wallich; Karen Wright
PUBLISHER John J. Moeling, Jr.
ADVERTISING DIRECTOR Robert F. Gregory
PRESIDENT AND CHIEF EXECUTIVE OFFICER Claus-Gerhard Firchow
CHAIRMAN OF THE BOARD Georg-Dieter von Holtzbrinck
CHAIRMAN EMERITUS Gerard Piel

**SUSCRIPCIONES**

Prensa Científica S. A.
Viladomat, 291 6º 1ª
08029 Barcelona (España)
Teléfonos 321 81 91 - 321 83 48

**Precios de suscripción, en pesetas:**

| | Un año | Dos años |
|---|---|---|
| España | 5500 | 10.000 |
| Extranjero | 6200 | 11.500 |

**Ejemplares sueltos:**

Ordinario: 500 pesetas
Extraordinario: 650 pesetas

- Todos los precios indicados incluyen el IVA, cuando es aplicable.
- En Canarias, Ceuta y Melilla los precios incluyen el transporte aéreo.
- El precio de los ejemplares atrasados es el mismo que el de los actuales.

**DISTRIBUCION**

**para España:**

MIDESA
Carretera de Irún, km. 13,350
(Variante de Fuencarral)
28049 Madrid Tel.652 42 00

**para los restantes países:**

Prensa Científica, S. A.
Viladomat, 291 6º 1ª - 08029 Barcelona
Teléfonos 321 81 91 - 321 83 48

**PUBLICIDAD**

Madrid: Gustavo Martínez Ovín
Menorca, 8, bajo, centro, izquierda.
28009 Madrid
Teléfonos 409 70 45 - 409 70 46

Cataluña:
Prensa Científica, S.A.
Viladomat, 291 6º 1ª
08029 Barcelona
Tels. 321 81 91 - 321 83 48






ISSN 0210-136X Dep. legal: B. 38.999-76

Fotocomposición: Tecfa. Línea Fotocomposición, S.A. Almogàvers, 189 - 08018 Barcelona
Fotocromos reproducidos por Tecfa, S.A. Almogàvers, 189 - 08018 Barcelona
Imprime Rotographik, S.A. Ctra. de Caldes, km 3,7 - Santa Perpètua de la Mogoda (Barcelona)

Printed in Spain - Impreso en España

Los espacios en gris
corresponden a publicidad
en la edición impresa

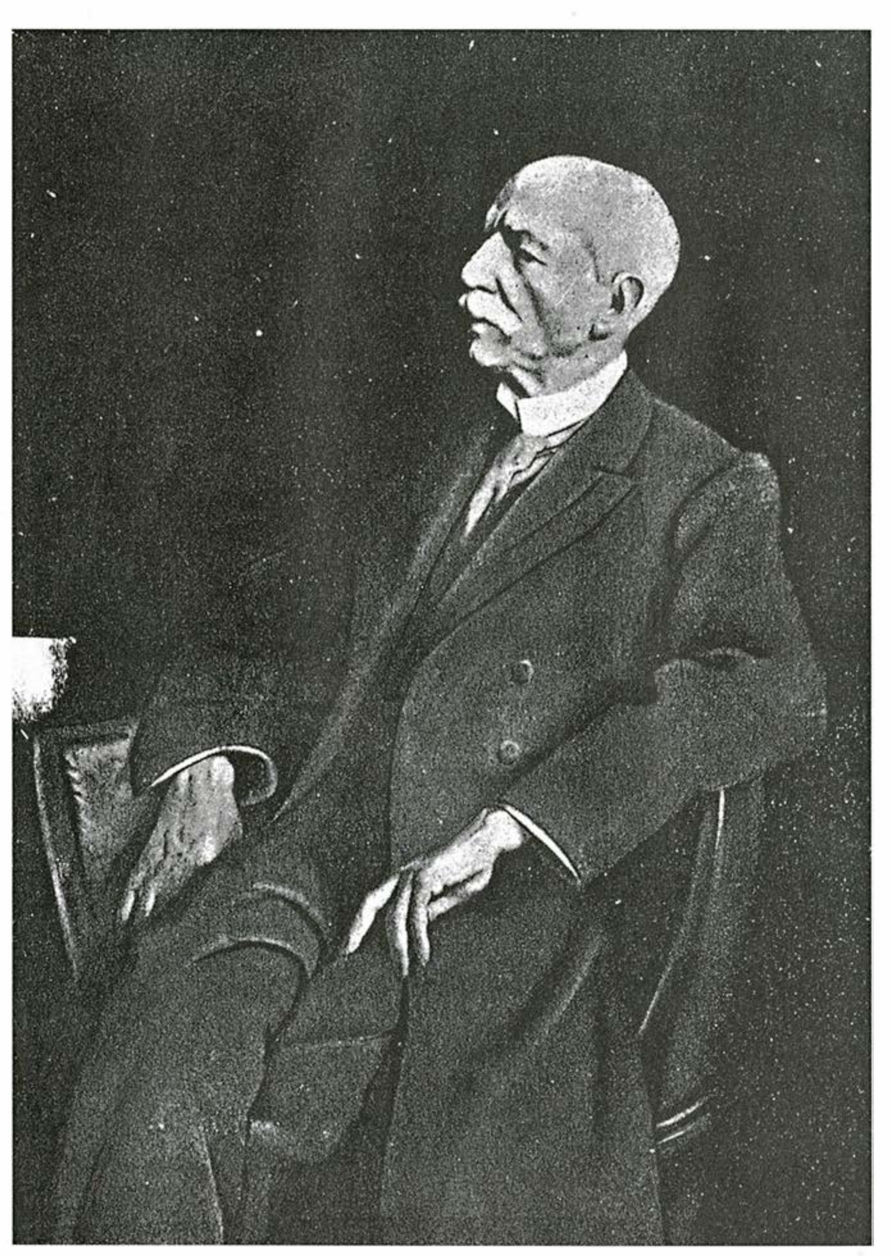
# Hace

José M.ª López Piñero

## *...ciento cincuenta años*

Manuel García presentó su Memoria sobre la voz humana ante la Academia de Ciencias, de París.

Nacido en Madrid el año 1805, Manuel García, cuyo verdadero nombre era Manuel Rodríguez Siches, fue hijo del tenor sevillano Manuel Rodríguez, llamado *el García*, y de la soprano Joaquina Siches. También sus hermanas María y Paulina se dedicaron a la ópera, consiguiendo la primera de ellas fama mundial con el nombre artístico de *la Malibrán*. Durante su juventud, Manuel García hizo algunos papeles de barítono en la compañía de su padre pero, convencido de la medianía de su voz, decidió en 1829 convertirse en profesor de canto.



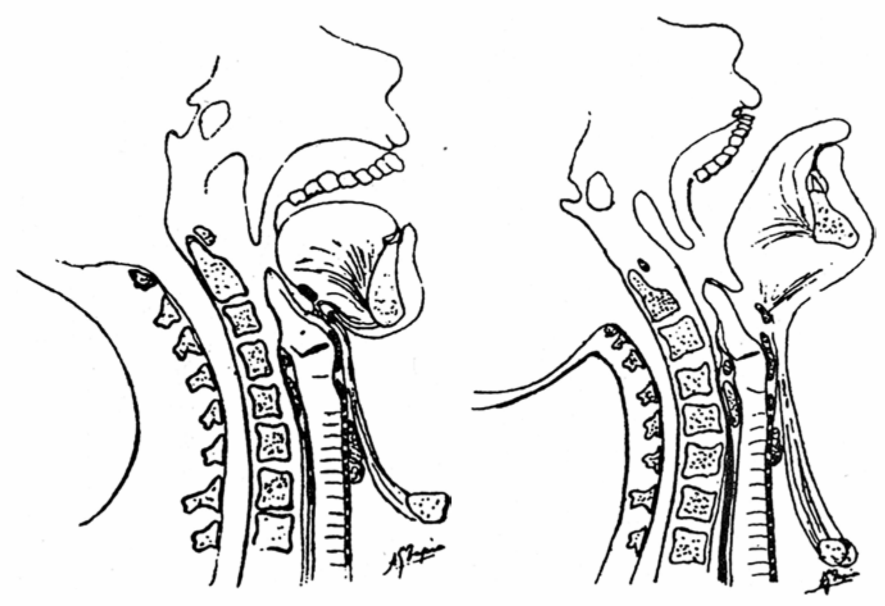
*1. Manuel García a los cien años. Retrato pintado por John Sargent que regalaron en 1905 al descubridor del laringoscopio sus discípulos y las asociaciones europeas de laringólogos.*

Se instaló en París, donde simultaneó sus estudios musicales con la investigación en torno a la laringe y el mecanismo de producción de la voz, tema para el que le resultó de gran utilidad trabajar en varios hospitales militares y conocer directamente los fenómenos que se presentan en las enfermedades de la garganta.

En el terreno de la enseñanza musical, que impartió primero como profesor del Conservatorio de París y desde 1855 en la Real Academia de Música de Londres, su publicación más importante fue un tratado sobre el arte del canto (1847) que alcanzó numerosas ediciones en varios idiomas. El principal fruto de sus investigaciones acerca de la producción de la voz fue el descubrimiento del laringoscopio, que comunicó a la Royal Society, de Londres, en la memoria titulada *Observaciones fisiológicas sobre la voz humana* (1855). Adoptado inmediatamente para la exploración laringológica por el vienés Ludwig Türck, al laringoscopio debió García su prestigio internacional en la comunidad médica, que se puso de manifiesto en el gran homenaje que, organizado por la Royal Medical and Chirurgial Society, de Londres, recibió al cumplir cien años en 1905, uno antes de su muerte.

Mucho menos recordada pero igualmente importante fue su *Memoria sobre la voz humana*, de la que un informe, firmado, entre otros, por el gran fisiólogo François Magendie, afirmó que contenía observaciones y descripciones de muchos hechos nuevos interesantes, "los cuales será preciso tener en cuenta, de ahora en adelante, en la teoría física de la voz humana". Reproduciremos a continuación un fragmento de dicha memoria relativo a la producción de los registros:

"Cuando se quiere hacer una honda inspiración, se fija la laringe todo lo más bajo posible, gracias a la relajación de todos los músculos extrínsecos, salvo los depresores. La laringe adquiere así su mayor capacidad. La faringe misma se dilata a causa del aire que se precipita a los pulmones.



*2. Esquemas del gran otorrinolaringólogo Antonio García Tapia incluidos en su libro sobre Manuel García (1905).*

Si en estas condiciones se espira con fuerza, todo el aire inspirado se escapa al instante, sin producir el menor sonido. Este hecho prueba que tal ampliación de la laringe no es de ninguna manera favorable a la fonación.



"Trátase de dar el sonido más grave del registro de pecho; al instante mismo se siente una ligera constricción interna y externa, y la laringe se fija un poco por encima de la posición precedente.

"Ningún esfuerzo es capaz de desalojar el aire ahora tan rápidamente como fue posible hacerlo en el caso anterior. Esta imposibilidad prueba la constricción interna. Para amortiguar la salida del aire, y obtener al mismo tiempo vibraciones, es preciso que la parte más estrechada del instrumento, y la más vibrátil a la vez, es decir, la glotis, se coloque a guisa de cuña sobre la tráquea, oponiendo así una resistencia al aire, y entrando en vibración por la acción de este último. La glotis es, pues, el punto donde se engendra el sonido más grave de la voz de pecho. (Todo conduce a creer que la glotis, que engendra el primer sonido grave, engendra igualmente el registro en toda su extensión.)

"La posición más baja de la laringe es la única que permite producir este sonido; en vano se intentaría obtenerle con una posición más elevada.

"Una nota cualquiera del registro de pecho puede siempre ser producida por un *coup de la glotte*, que de antemano está cerrada. El aire acumúlase allí, y de repente se abre la glotis para producir las vibraciones. Esta explosión glótica puede ser áfona.

"La vibración del músculo llamado tiro-aritenoideo ha sido ya demostrada por los experimentos de Magendie en perros vivos. Estos experimentos permitieron ver las vibraciones de las cuerdas, poniendo la glotis al descubierto. Las vibraciones subsisten aunque se corten los labios superiores; pero cesan, y con ellas la voz, cuando se cortan los nervios laríngeos o la glotis inferior. Estos hechos, unidos a las observaciones a que han dado lugar la insuflación de laringes en cadáveres, y las fístulas traqueales en sujetos vivos, hacen pensar que, no solamente vibran los labios inferiores durante la producción del registro que nos ocupa, sino que, además, su aproximación gradual, su tensión proporcional acompañada de la disminución progresiva de la capacidad de la laringe, son los principios que presiden a la formación del registro de pecho, desde los sonidos más graves a los más agudos.

"Sarvat ha dicho el primero que, para que los sonidos de este registro reuniesen todas sus cualidades, es preciso que la tensión de la parte extensible de pared del tubo vocal esté en relación conveniente con la de las paredes de los ventrículos, así como con la de los ligamentos superiores e inferiores, y que las dimensiones de los orificios a través de los cuales se escape el aire puedan asimismo variar y amoldarse convenientemente para dar el mejor resultado posible. Los repliegues de la mucosa, que flotan en medio del aire que resuena alrededor de ellos, de cuyo movimiento participan, deben tener por principal objeto, en razón de su tensión variable, modificar, ya sea súbita, ya gradualmente, el número de vibraciones del aire.

"Magendie atribuye a los ventrículos otra función. Según este sabio profesor, los ventrículos tienen por objeto aislar las cuerdas inferiores de las paredes de la laringe. La observación que nosotros hemos hecho de la laringe de un buey parece venir en apoyo de esta opinión. Estando desprovista la laringe del buey de ventrículos, no pueden aislarse las cuerdas vocales, sino por medio de las apófisis internas de los cartílagos aritenoides, que se encuentran muy prolongadas. Cuando las dos apófisis se ponen en contacto, arrastran la una hacia la otra las cuerdas vocales, que así se colocan delante de la columna de aire por su borde libre, y se adelgazan cada vez más. Separadas así las paredes de la laringe, los labios de la glotis tienen, tanto por encima como por debajo, todo el espacio necesario para efectuar vibraciones extensas.

"En cuanto al papel de los labios superiores de la glotis, no está bien determinado. Es evidente que formando el techo de los ventrículos no

pueden permanecer extraños a la producción del registro que nos ocupa; pero, ¿cuál es su papel? ¿Es pasivo, es activo?

"La glotis inferior es la más importante de las dos, por ser el punto donde se engendra la serie de sonidos que se ha convenido en llamar *registro de pecho*. Nos parece que no debe conservarse por más tiempo esta denominación ni la de registro laríngeo; la primera da una idea falsa del sitio donde nace la serie de sonidos de que se trata; la segunda no determina el punto preciso donde nacen las vibraciones. ¿No valdría la pena sustituirlos con un nombre más preciso?

"Cuando se eleva gradualmente la voz, del registro de falsete al registro de cabeza, empleando durante el primero el *timbre claro* y durante el segundo el *timbre obscuro*, el que escucha nota mayor claridad y pureza en la voz a partir de las notas $re_3$ $mi_3$ y el que canta experimenta una sensación de repercusión bastante pronunciada, hacia la apófisis basilar. Estos efectos han hecho creer que las notas $re_3$ y $mi_3$ eran la separación de los dos registros, el uno inferior, de falsete, y el otro superior, de cabeza. Nosotros vamos a combatir esta creencia, puesto que tal separación no es el resultado de un mecanismo diferente de la laringe, sino un efecto de *timbre*, como ya hemos indicado.

"La diferencia de efecto, la precisión del límite que en el caso predicho separa estas dos modalidades de la voz humana, desaparece desde el momento en que se pasa de una a otra sin variar de tiempo.

"A esta consideraciones podemos añadir que, si estas dos modalidades de voz que nos ocupan constituyeran dos registros diferentes, la laringe habría de comportarse siguiendo las leyes que rigen a los registros de naturaleza y mecanismo diferentes. Es decir, que para la producción de las notas más bajas de cada uno, la laringe adoptará su posición más baja, según hemos visto, entre el registro de pecho y el de falsete, y además, sería preciso que, entre los sonidos propios de la voz de falsete y los de cabeza, hubiese una serie de notas comunes, o, por el contrario, hubiera una laguna, como veremos existe entre los registros de *pecho* y *contrabajo*.

"Muy lejos de esto, las notas de la voz de falsete y las de la voz de cabeza, se continúan y forman una mutua prolongación, durante la cual la laringe se comporta como lo exige la sucesión de sonidos de un solo registro; de donde efectivamente se infiere que sólo existe uno".




**1. DEL FRUTO DE LA ADORMIDERA** (*Papaver somniferum*) **se obtiene el opio, uno de cuyos elementos componentes es la morfina. Esta droga se viene usando en terapia médica desde hace más de dos mil años. (La pintura está tomada del célebre ilustrador francés del siglo XVIII Rédouté.)**

# La tragedia del dolor innecesario

*En contra de lo que la gente cree, la morfina tomada sólo para controlar el dolor no provoca adicción. Pero se sigue administrando a los pacientes dosis del todo insuficientes, dejándoles sufrir una agonía innecesaria*

Ronald Melzack

Como dijo en cierta ocasión Albert Schweitzer, "el dolor es para la humanidad un tirano más terrible que la misma muerte". El sufrimiento prolongado destruye la calidad de la vida y puede quitar las ganas de existir, llevando a algunas personas al suicidio. Sus efectos físicos son también muy destructores: un dolor fuerte y persistente puede impedir el sueño y quitar el apetito, produciendo con ello el agotamiento y reduciendo la disponibilidad de nutrientes para los órganos. El dolor podría así obstaculizar la recuperación de una enfermedad o de una lesión y, en enfermos muy debilitados o de edad avanzada, llegar, tal vez, a constituir la última diferencia entre la vida y la muerte.

Por desgracia, hay algunas clases de dolor que los tratamientos conocidos no pueden aliviar. Que en esos casos la asistencia médica logra bien poco, resulta muy angustiante para todos los implicados, pero es sin duda comprensible. Menos comprensible es que muchas personas sufran, no porque se carezca de medios para aliviar el dolor, sino porque a menudo los médicos son reacios a prescribir morfina. El más seguro y eficaz analgésico (o antidolórico) de cuantos se conocen para combatir el dolor agudo y persistente es la morfina. Pero como este fármaco provoca adicción en algunas personas, se la suele administrar, si es que se la administra, sólo en minúsculas dosis.

De hecho, la preocupación por la adicción ha motivado que en muchos países de Europa y en otras partes la ley prohíba virtualmente todos los usos de la morfina y sustancias afines, incluidas sus aplicaciones médicas. Hasta allí donde la terapia con morfina es legal, como ocurre en la Gran Bretaña y en EE.UU., muchos doctores y enfermeros, temerosos de convertir a sus pacientes en adictos, les suministran dosis insuficientes o demasiado espaciadas como para que puedan así controlar el dolor.

Sin embargo, lo cierto es que entre los pacientes que toman morfina para combatir el dolor raras veces se dan casos de adicción –caracterizada por una ansia psicológica de recibir una determinada sustancia, y, si ésta se suprime de repente, por el desarrollo de unos síntomas de su carencia (sudoraciones, dolores musculares y náuseas). La adicción parece producirse tan sólo en un sector de los usuarios, en aquellos que toman la morfina para disfrutar de sus efectos psíquicos, tales como su poder euforizante y aliviador de tensiones.

Además, en los pacientes que toman morfina contra el dolor no se desarrolla la rápida tolerancia física a la droga, que es, a menudo, un signo de la adicción. Muchos individuos propensos a la adicción necesitan en seguida dosis cada vez más fuertes para lograr el deseado cambio de estado de ánimo, mientras que los pacientes que toman la droga con fines analgésicos no requieren, para conseguir el alivio, fuertes incrementos en las dosis. Inicialmente quizá se origine en ellos alguna tolerancia, pero la dosis necesaria suele aumentarse gradualmente y, después, se estabiliza.

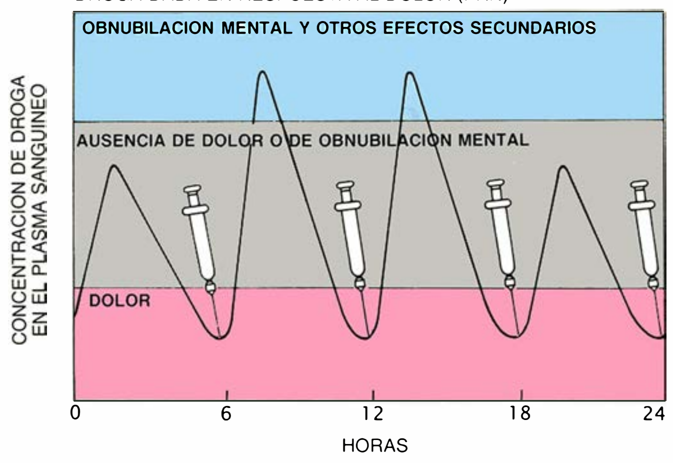
En modo alguno sugiero que la morfina se recete de manera indiscriminada. Lo que sí pretendo es urgir a los legisladores, a los encargados de hacer cumplir las normas sanitarias y a todo el personal que atiende a los enfermos que disciernan entre el adicto que ansía la morfina por sus propiedades euforizantes y el paciente psicológicamente sano que toma la droga con el fin exclusivo de aliviar un dolor.

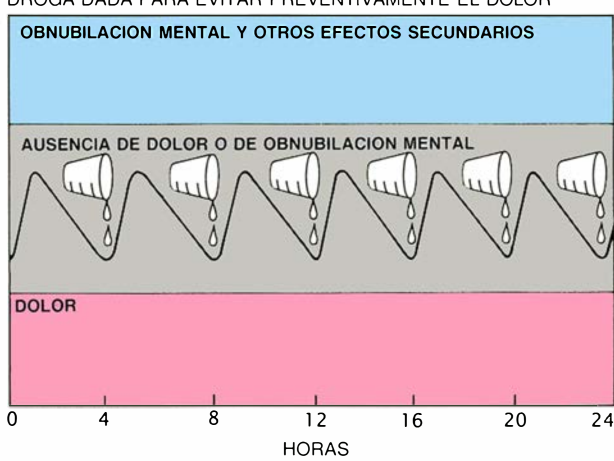
La morfina es uno de los elementos constitutivos del opio, agente terapéutico utilizado desde hace más de 2000 años, al menos desde los tiempos de la antigua Roma. Obtiénese el opio extrayendo un jugo lechoso de la cápsula inmadura de la adormidera (*Papaver somniferum*, planta que abunda en muchos países del Oriente Medio) y secando a continuación ese exudado para formar una goma. Esta goma –el opio– puede ingerirse tal cual o añadirse a un brebaje.

Allá por el siglo XVI el opio era objeto de comercio entre Europa y el Oriente. Por aquel entonces, una mixtura que contenía opio y que recibió el nombre de láudano llegó a ser en Europa un remedio popular para casi todo tipo de indisposiciones. Andando el tiempo, el fumar opio junto con el tabaco se convirtió en una manera no menos popular de disfrutar esta droga.

A comienzos del siglo XIX, Friedrich W.A. Sertürner, joven farmacéutico alemán, aisló la morfina del opio y la identificó como el principal ingrediente activo. Después de la ob-


RONALD MELZACK, que estudia la neurofisiología del dolor desde hace 35 años, es profesor en la cátedra E.P. Taylor de psicología en la Universidad McGill y director de investigación en el Hospital General de Montreal. Tras doctorarse en psicología por la McGill en 1954 y disfrutar de becas en los EE.UU. y en el extranjero, entró a formar parte del claustro del Instituto de Tecnología de Massachusetts en 1959. Allí trabajó con Patrick D. Wall, colaboración que conduciría a ambos a la publicación, en 1965, de su actualmente famosa teoría sobre el "control de entrada" del dolor. Regresó a McGill en 1963.






**2. PROGRAMA ESTANDAR para mitigar, de forma continua, con morfina, un dolor constante (*izquierda*). Consiste en prescribir que las inyecciones se administren 'pro re nata' (PRN), o sea, "según necesidad", cuando acucie de nuevo el dolor. Esto significa, en la práctica, que las inyecciones se ponen sólo como respuesta al dolor, y que, si éste retorna antes de que hayan pasado de cuatro a seis horas, el paciente deberá a menudo esperar hasta la próxima dosis. Cuando recibe la inyección siguiente, el dolor puede que sea ya tan terrible que haga falta para calmarlo una dosis muy grande, que le producirá obnubilación mental y otros efectos secundarios tales como náuseas. Un planteamiento más avisado (*derecha*) procura la evitación real del dolor, con lo que contribuye a que el enfermo pierda el miedo a la recurrencia de su suplicio. La morfina se administra oralmente (en una dosis calculada según las necesidades de cada paciente) cada cuatro horas, o incluso con mayor frecuencia si de este modo resulta más efectiva. Las dosis son frecuentes, y pueden así ser de ordinario relativamente bajas, lo cual reduce la incidencia de efectos secundarios.**

tención de morfina, en 1832 se separó otro opiáceo o derivado del opio: la codeína.

A mediados del siglo XIX, la introducción de la aguja hipodérmica permitió administrar por inyección gran cantidad de drogas. Este adelanto técnico, que facilitaba el rápido efecto de la droga, motivó que en muchos países del mundo entero se tendiese a prescribir en seguida inyecciones de morfina para mitigar los dolores intensos. Al mismo tiempo, cada vez fueron más los individuos que empezaron a tomar morfina por sus efectos emocionales, y aumentó el número de los adictos.

Con posterioridad comenzó la búsqueda de drogas que tuviesen las propiedades analgésicas de la morfina, pero no crearan hábito. Esta investigación dio por resultado la producción de la heroína, compuesto sintético cuya actividad era similar a la de la morfina, aunque, por desgracia, provocaba también drogodependencia. Entraron luego en escena otros opioides (sustancias químicas que actúan de modo parecido al del opio), entre ellos la metadona y la meperidina (Demerol). Como los opiáceos, muchos de los opioides alivian el dolor, inducen cambios del estado de ánimo y causan algún grado de adicción.

Inevitablemente, el creciente abuso de los narcóticos (me refiero a opiáceos y opioides) y de otras drogas que alteran el estado anímico incitó a los gobiernos de la mayoría de los países a imponer reglamentos antidroga. Y, a la vez, se fue extendiendo la costumbre de extremar la cautela en todo lo concerniente a la administración de drogas para calmar el dolor.

Hoy día la terapia analgésica mediante morfina suele restringirse a dos clases de pacientes: se prescribe durante períodos bastante cortos para aquellos pacientes hospitalizados que sufren dolores después de una intervención quirúrgica, o bien durante períodos, que pueden ser más largos, para aliviar los dolores de las quemaduras o de los cánceres incurables.

En muchos hospitales es común que a la prescripción de morfina se añada las siglas "PRN" (que significa *pro re nata* o "según necesidad"). Con esto se quiere indicar que sólo se habrá de administrar la droga después de la reaparición del dolor. Lo usual es que se administre por inyección intramuscular o subcutánea.

Con frecuencia, la prescripción PRN provoca una confrontación entre el paciente y el personal asistente, que espera que el efecto analgésico de la morfina dure de cuatro a seis horas. El paciente, que ha vuelto a sentir el dolor antes de lo esperado, sufre lo indecible y suplica que se le ponga la siguiente inyección. Pero el asistente, temiendo causar la adicción, rehúsa complacerle. Cuando se procede, al fin, a mitigar el dolor, puede que éste sea tan intenso que se requiera una gran dosis, lo cual aumenta la probabilidad de que se produzcan efectos secundarios, tales como la obnubilación mental y náuseas. Sobre todo cuando un paciente sufre una enfermedad terminal carece de sentido preocuparse por la adicción y es cruel demorarle el alivio.

Hay otro modo más humano de tratar el dolor, un modo que va obteniendo paulatinamente mayor aceptación. Consiste en administrar las dosis con regularidad, a tenor de un programa especialmente diseñado para evitar la recurrencia del dolor. De esta forma, se impide en todo momento el dolor, pues el paciente no ha de esperar a sentir nuevas molestias para recibir la dosis siguiente.

Este nuevo enfoque preventivo fue cobrando auge a partir de la tarea iniciada hace unos 20 años por la doctora inglesa Cicely M. Saunders, quien fundó el primer centro moderno dedicado a la asistencia de enfermos de cáncer u otras patologías en fase terminal: el St. Christopher's Hospice de Londres. Exhortaba Saunders a los médicos que atendían a los enfermos terminales a que se enfrentaran con la realidad y procurasen más bien mitigar los dolores de sus pacientes y evitarles las náuseas y demás molestias, en vez de tratar inútilmente de curar sus males. Creía ella que los días o semanas últimos de la vida de una persona deberían ser un tiempo de reposo y de bienestar, pasado lo más agradablemente posible en compañía de los familiares y los amigos.

Con esta finalidad, Saunders prescribía la mezcla Brompton, variante de un analgésico líquido que había sido empleado para los cánceres avanzados en varios hospitales londinenses, entre ellos el Hospital Brompton de Enfermedades del Tórax, desde finales del siglo XIX. Este remedio (mezcla de morfina, cocaína, cloroformo diluido en agua, alcohol y

un jarabe aromatizante) se había arrinconado en favor de la morfina inyectable, pero Saunders se dio cuenta de que un fármaco tomado por vía oral permitiría a muchos pacientes pasar algunos de sus últimos días en el hogar; bastaría la visita de una enfermera que regulara las tomas del producto para que no sufriesen ningún dolor.

Se averiguó más tarde que el único ingrediente importante de la mezcla Brompton era la morfina, de modo que, en la actualidad, los pacientes tratados con el enfoque preventivo respecto del dolor sólo toman morfina, ya sea en forma de pastillas, ya mezclada con bebidas. Normalmente se administra una dosis inicial de 10 miligramos, que se repite cada cuatro horas. Después, transcurridos unos días, o semanas según los casos, se van ajustando las dosis y su periodicidad hasta establecer un régimen de mantenimiento a raya del dolor, con un horario calculado para que no se produzcan obnubilaciones ni otros efectos secundarios.

Un tratamiento que otorgue especial atención a la prevención del dolor es de aconsejar en los casos de enfermos de cáncer. El dolor y el miedo a padecerlo constituyen quizá las fuentes principales de su sufrimiento. En los estadios iniciales de la enfermedad, en alrededor del 50 por ciento de los pacientes el dolor procede del propio cáncer o de las medidas puestas en práctica para detener su extensión. Cuando el cáncer ha llegado a sus fases finales, aproximadamente un 70 por ciento de los pacientes se quejan de dolores que tienden a ser intensos y persistentes.

Del 80 al 90 por ciento de los enfermos de cáncer tratados con un enfoque preventivo del dolor consiguen un alivio satisfactorio y declaran que su malestar se les hace bastante soportable o, con más frecuencia, que ha desaparecido del todo. Y la mitad, más o menos, de los restantes obtienen alivio si se les aplican además otras terapias. Esta proporción de buenos resultados es digna de nota, habida cuenta de lo destructivo que es el cáncer y de los fuertes dolores que lo acompañan.

Los tratamientos se perfeccionan sin cesar. Hay ahora cápsulas especiales que liberan la morfina lentamente, de modo que sólo es necesario tomarlas unas pocas veces al día. Dispónese también de bombas portátiles, controladas electrónicamente, que liberan bajo la piel infusiones continuas del medicamento.

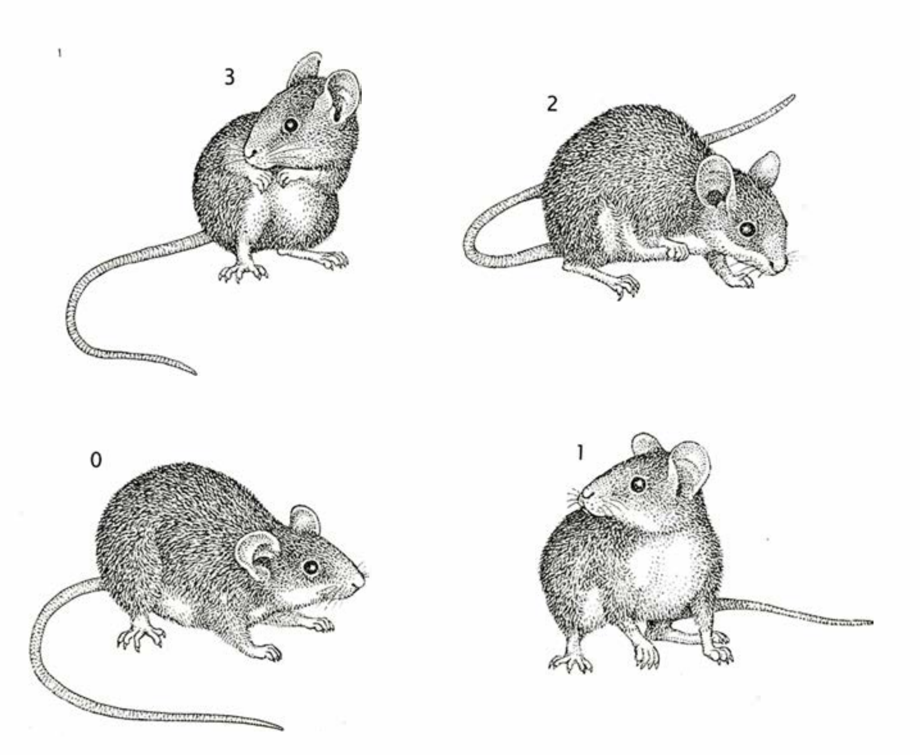
Se han reunido ya suficientes pruebas para demostrar que el tradicional enfoque PRN, basado en el miedo a la adicción, tiene escaso sentido. Numerosos estudios de pacientes que sufren y son tratados principalmente con narcóticos –a saber, los enfermos de cáncer, las víctimas de quemaduras y los hospitalizados por intervenciones quirúrgicas– han evidenciado que los sujetos en quienes se desarrolla una rápida y señalada tolerancia a los narcóticos, y dependencia de los mismos, suelen ser aquellos que tienen ya un historial de perturbaciones psíquicas o de abuso de la sustancia administrada.

Consideremos en primer lugar el problema de la tolerancia notoria, que no sólo es un indicio de posible adicción, sino también tema de interés médico por su propio derecho, puesto que el peligro de efectos secundarios aumenta con la dosis. Por ejemplo, la aplicación de morfina en cantidades excesivas puede acarrear el coma y obstaculizar gravemente las funciones respiratorias.

Robert G. Twycross, actualmente en el Hospital Churchill de Oxford, ha demostrado que es bastante pequeña la tolerancia que se desarrolla en pacientes cancerosos que toman varias veces al día, durante largos períodos, dosis de heroína ajustadas a sus necesidades específicas. Los pacientes mostraban cierta tolerancia a la droga en la fase inicial, por lo que hubo que ir aumentando gradualmente las dosis durante las 12 primeras semanas, pero el alivio del dolor se conseguía sin que se produjesen efectos secundarios graves. A continuación, las dosis se mantenían casi estables durante meses.

Balfour M. Mount, uno de mis colegas en la Universidad McGill, y yo mismo obtuvimos recientemente resultados similares al estudiar la tolerancia a la morfina en pacientes que llevaban más de un mes en la Unidad de Atención Paliativa del Hospital Reina Victoria de Montreal. (Esta unidad, fundada por Mount, fue el primer servicio de cuidados paliativos en un gran hospital general.) Los enfermos de nuestro estudio, a los que se les administraba la droga por vía oral, respondieron a un cuestionario que, para la evaluación del dolor, compuse con Warren S. Torgerson, de la Universidad Johns Hopkins. Definimos una escala de grados de intensidad del dolor, que iba desde su ausencia (0), pasando por el dolor suave (1), el ya más molesto (2), el angustiante (3), el horrible (4), hasta el absolutamente insoportable y enloquecedor (5).

Alrededor de un 5 por ciento de los pacientes tenían incesantemente do-




**3. CON EL TEST DE LA FORMALINA se miden los efectos analgésicos del dolor 'tónico' o persistente. Una disolución de formaldehído en solución salina se inyecta subcutáneamente en la patita de una rata, provocándole así un dolor que dura unos 90 minutos. El animal se lame repetidamente la pata, lo que es señal de moderado dolor (un dolor de grado 3). Luego, al cabo de unos momentos, mantiene su pata levantada (grado 2), la apoya después cautelosamente en el suelo (grado 1 del dolor) y, por último, camina ya normalmente (grado 0). En este test, las ratas tratadas con morfina muestran débil tolerancia al efecto analgésico de la droga: no requieren dosis cada vez mayores para lograr alivio. Tal observación concuerda con los resultados clínicos que demuestran que los pacientes que toman morfina para aliviar el dolor persistente ni adquieren tolerancia ni se hacen adictos.**

lores de altas gradaciones (del 3 para arriba). El 85 por ciento restante habían logrado un excelente control del dolor sin necesidad de rápidos aumentos de las dosis de morfina.

La agudización del dolor, que suele ser indicio de un avance de la enfermedad posterior al establecimiento de un programa de mantenimiento, era la razón más común de incremento de dosis. Los pacientes que notaban que sus molestias habían remitido –ya fuese espontáneamente, ya gracias al tratamiento, como, por ejemplo, tras la reducción de un tumor mediante radioterapia– solían requerir menos medicación.

John F. Scott, del Hospital Elizabeth Bruyere, en Ottawa, encontró pocos indicios de adicción al analizar numerosos estudios en los que se examinaban los síntomas de abstinencia en pacientes de clínicas especializadas en el tratamiento del cáncer. Su informe señala que, "si un enfermo de cáncer no requiere ya un narcótico para controlar el dolor, una reducción gradual de la dosis evitará todos los síntomas de abstinencia, aunque éstos suelen ser mínimos o nulos, aun después de interrumpirse las dosis bruscamente". Toda dependencia física es superada, por lo general, sin dificultad cuando se van reduciendo las dosis en un período de unos cuantos días.

Estudios de pacientes que recibían narcóticos mientras estaban hospitalizados tampoco aducen pruebas sólidas de procesos de adicción. En una investigación exhaustiva, Jane B. Porter y Hershel Jick, del Hospital Clínico de la Universidad de Boston, analizaron 11.882 historias de pacientes que recibieron narcóticos para aliviar el dolor derivado de diversos problemas médicos; no se encontró ninguna anotación de drogodependencia. Sólo cuatro de los pacientes abusaron después de las drogas, y exclusivamente en uno de ellos los niveles de consumo fueron graves.

No menos persuasivos son los resultados del estudio de más de 10.000 quemados en distintos grados. A estos individuos, estudiados por Samuel W. Perry, del Hospital de Nueva York, y George Heidrich, de la Universidad de Wisconsin en Madison, se les tuvo que practicar la desbridación, procedimiento dolorosísimo en el que se separa el tejido muerto de la piel quemada. La mayoría de ellos recibieron inyecciones de narcóticos durante semanas o incluso meses. Sin embargo, ni un solo caso de posterior adicción pudo atribuirse a los narcóticos que se administraron para aliviar el dolor durante su hospitalización. Aunque 22 pacientes abusaron del consumo de drogas después de ser dados ya de alta, todos ellos contaban con un historial de drogadicción.

Más pruebas de que los narcóticos pueden ser administrados para combatir el dolor sin que causen adicción nos las proporcionan estudios sobre "analgesia en pacientes controlados", en pacientes de intervenciones quirúrgicas y en internos quemados. En esos estudios, los pacientes pulsan un botón para accionar una bomba controlada electrónicamente y situada junto a su lecho, mediante la cual se autoadministran pequeñas dosis de morfina (con una intubación intravenosa). Cuando se introdujeron tales dispositivos se temía, y no poco, que los pacientes abusaran de la droga. Pero no tardó en patentizarse que los pacientes mantienen sus dosis a un nivel prudencial y las aminoran cuando disminuye el dolor.

Los estudios que exploran la analgesia producida por la morfina están ayudando a comprender la causa del carácter improbable de que, en los pacientes que sólo toman la droga para aliviar el dolor, se desarrolle una rápida tolerancia y de que se tornen adictos. Basándonos en tales estudios, mi antiguo discípulo Frances V. Abbott y yo propusimos en 1981 la teoría de que, probablemente, la morfina actúa sobre dos distintos sistemas de señalización del dolor pertenecientes al sistema nervioso central, y de que uno de ellos –el que da origen al tipo de dolores que se suele tratar con la morfina– no desarrolla de hecho mucha tolerancia a la droga.

Nuestra propuesta dimanó de mis esfuerzos por realizar en animales una serie de pruebas encaminadas a determinar con precisión la eficacia de las drogas analgésicas para el tipo de dolores que más a menudo requieren el uso de narcóticos en los pacientes humanos: los dolores prolongados o "tónicos", que perduran mucho después de sufrida la lesión. Este es el tipo de dolores que suelen atormentar crónicamente a los enfermos de cáncer. El dolor que se siente inmediatamente después de una lesión se califica de "fásico", dolor breve y de una intensidad que aumenta rápidamente para decaer también en seguida. (Un dolor fásico es, por ejemplo, el que sentimos en el instante de hacernos un corte en un dedo.) Al fásico suele seguirle el de tipo tónico.

Durante muchos años, los investigadores interesados en medir los efectos analgésicos de las drogas han sometido a ratas a lo que se conoce como el "test del tirón de cola". Tras inyectarle a una rata la droga en cuestión, se le mete la cola en agua muy caliente; el tiempo que transcurre entre la inmersión y el momento en que la rata saca de un tirón su rabo del agua es medido como un índice de dolor. Cuando se examinó con este test la eficacia de la morfina, los investigadores hallaron pruebas de una considerable tolerancia: los animales requerían dosis cada vez mayores para mantener la cola dentro del agua durante un tiempo dado. Se interpretó que tales resultados indicaban que los pacientes humanos con dolores se harían rápidamente tolerantes a la morfina y, así, se convertirían en adictos.

Hay, empero, una seria objeción respecto al test del tirón de cola. Y es que semejante procedimiento origina un dolor fásico, de repentino auge, que no es precisamente el tipo de dolor para el que suele prescribirse la morfina. Con el fin de obtener mayor información acerca de los efectos de los analgésicos sobre el dolor tónico, persistente, en los humanos, John O'Keefe, David Dubuisson y Stephen G. Dennis, que eran entonces alumnos míos, desarrollaron el test de la formalina. Una pequeña cantidad de formalina –formaldehído diluido en una solución salina– se inyecta hipodérmicamente a una rata en una de sus patas delanteras. Si no se le propina al animal ningún analgésico, la formalina le produce un moderado dolor que dura unos 90 minutos, como lo ponen de manifiesto la tendencia del animal a lamerse la pata y su renuencia a cargar peso sobre ella. Si una droga le calma el dolor, el animal tarda menos en apoyarse sobre la pata.

Con el test de la formalina (para el dolor tónico), Abbott y yo (luego, en McGill, se uniría Keith B.J. Franklin a nuestro equipo) averiguamos que en las ratas se desarrollaba una débil tolerancia a la analgesia producida por sucesivas inyecciones de morfina. La explicación más lógica de los diferentes grados de tolerancia detectados en los tests del tirón de cola y de la formalina era que el dolor fásico y el tónico son causados por dos sitemas neurales distintos que tienen diferente tolerancia a la morfina.

Otra serie de pruebas contribuyeron a corroborar esta idea. Por ejemplo, Dennis y yo examinamos el influjo que sobre el dolor ejercen varias drogas que modifican la acción de la morfina (o que alteran el dolor por sí mismas), tanto en el test del tirón de cola como en el de la formalina. Los resultados fueron sorprendentes: efectos de la droga observados en uno de los tests faltaban o eran del todo inversos en el otro. Por ejemplo, drogas que reducían la analgesia de la

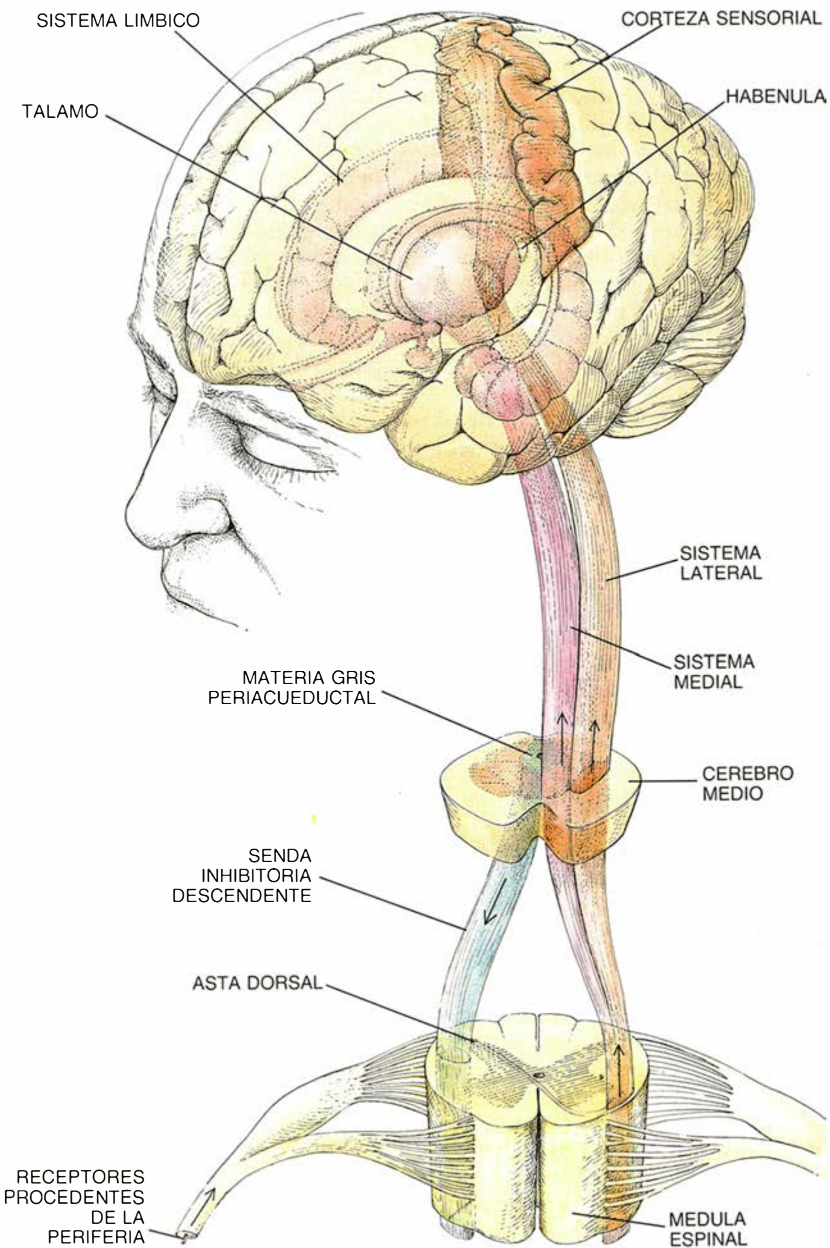
morfina en un test o no ejercían ningún efecto o aumentaban esa analgesia en el otro test. Si los sistemas neurales que responden al dolor fásico y al tónico fuesen el mismo, los efectos de las drogas sobre la actividad de la morfina deberían ser idénticos en ambos tests.

Mis colegas y yo creemos saber ahora cuáles de entre las muchas sendas neurales que hay en la médula espinal y en el cerebro constituyen los dos sistemas señalizadores del dolor que son sensibles a la morfina. Sabemos también algo sobre su funcionamiento y sobre su sensibilidad a la morfina. En ambos sistemas, la información relativa al dolor es enviada a las astas dorsales (regiones en forma de alas) de la médula espinal por neuronas periféricas que proceden de la piel y de otros tejidos del cuerpo [*véase la figura 4*]. Las neuronas ascendentes, que se originan en las astas dorsales, transmiten luego hacia arriba las señales del dolor, las cuales, recorriendo la médula espinal, van a parar a diversas partes del cerebro.

El sistema señalizador del dolor que mis colegas y yo pensamos que más relación tiene con el dolor repentino, fásico, recibe el nombre de sistema lateral. Se le llama así por el simple hecho de que sus prolongaciones, que llegan hasta el córtex sensorial, atraviesan el tronco encefálico en posición lateral respecto a la zona central del mismo. Al sistema que probablemente corresponde al dolor persistente, tónico, se le denomina sistema medial, porque sus prolongaciones pasan por la parte central del tronco encefálico.

Entre las propiedades más destacadas del sistema lateral se cuentan la rápida conducción de los impulsos y una organización que registra la posición relativa de los puntos del cuerpo. Estas propiedades capacitarían al sistema para originar dolor repentino y agudo en un punto del cuerpo prontamente identificado. Kenneth L. Casey, de la Universidad de Michigan en Ann Arbor, y yo hemos propuesto que las prolongaciones laterales son también responsables en gran parte de las cualidades sensoriales del dolor, como por ejemplo de su carácter punzante o ardiente.

La actividad del sistema lateral se amortigua, al parecer, bastante de-




**4. DOS SISTEMAS de neuronas son los responsables del dolor: un sistema medial (*color rosado*), que pasa a través de la zona central del tronco encefálico, y otro sistema lateral (*color naranja*). Ambos son bilaterales, constan de varios tractos y transmiten a los centros superiores las señales dolorosas que entran por las astas dorsales de la médula espinal. Se cree que el sistema medial es el principal responsable del dolor persistente (tónico). Envía señales al sistema límbico del cerebro, que influye en las emociones; en virtud de ello, se piensa también que da origen al componente afectivo del dolor (al que se alude con calificaciones tales como "espantoso" o "cruel"). Parece ser que el sistema lateral entra en acción sobre todo durante el dolor fásico, dolor que es repentino y agudo. Este envía señales al córtex sensorial, por lo que se supone que da origen a sensaciones tales como los calambres o los picores. La morfina puede inhibir ambos sistemas, pero el medial (dolor tónico) desarrolla mucha menos tolerancia a los efectos analgésicos de la droga que el lateral (dolor fásico) –lo que quizás explique por qué los pacientes que toman morfina para aliviar el dolor persistente (o tónico) no muestran de hecho gran tolerancia a esta droga. La morfina produce analgesia, en parte, inhibiendo la afluencia de las señales dolorosas desde los nervios periféricos a las sendas ascendentes; actúa de modo directo en las astas dorsales y activa también un sistema inhibidor descendente (*color azul*) que se origina en el cerebro medio. La morfina actúa, asimismo, en zonas situadas encima de la materia gris periacueductal, incluidos el sistema límbico y la habénula, asociada a los sistemas medial y límbico. Tal actividad parece contribuir al efecto analgésico.**

prisa, lo que explicaría por qué a menudo el dolor fásico remite o se calma con prontitud. La inhibición se lleva a cabo por un sistema de neuronas que tiene su origen en la materia gris periacueductal, situada en la parte del tronco encefálico que se conoce como cerebro medio. Este sistema descendente envía señales hacia abajo hasta las astas dorsales, donde inhibe la transmisión de las señales dolorosas desde los nervios periféricos hasta los tractos ascendentes. Tras un golpe o una lesión, parece activarse por los opioides (endorfinas y encefalinas) del propio cuerpo.

Si, como sugerimos, el sistema lateral transmite las señales que suscitan el dolor repentino, fásico, entonces no es sorprendente que el sistema esté por naturaleza sometido a una fuerte inhibición. El repentino dolor provocado por una herida o un golpe recientes podría abrumar o embotar al animal hasta el punto de impedirle defenderse luchando, correr a ocultarse o huir de su predador en una emergencia.

El otro sistema señalizador del dolor –el sistema medial– difiere del lateral en muchos aspectos. Por ejemplo, algunas de sus prolongaciones envían impulsos al sistema límbico, que comprende las regiones subcorticales del cerebro implicadas en la motivación y en las emociones. Por eso pensamos que el sistema medial controla el componente emocional del dolor, produciendo cualidades que podrían describirse con adjetivos como desgarrador, espantoso, atroz y similares. El sistema influye también en las actitudes que toma uno en respuesta a tales sensaciones.

Como el sistema medial transporta las señales con relativa lentitud a través de muchas neuronas pequeñas, no es muy apropiado para proporcionar información precisa durante las emergencias. Es apto, más bien, para producir sensaciones difusas de desagrado por algún tiempo con posterioridad a la ocurrencia del perjuicio, daño o lesión. Tales sensaciones ayudarían a asegurar que, habiendo sobrevivido a un peligro inmediato, el individuo herido en el lance se sienta muy abatido y, así, permanezca inactivo el tiempo suficiente para sanar.

¿Dónde ejerce la morfina sus efectos? Tanto en el sistema lateral (o del dolor fásico) como en el medial (dolor tónico), la morfina ejerce claramente algún efecto en las astas dorsales. Se sabe también que la morfina puede activar el sistema inhibidor descendente que se origina en la materia gris periacueductal. Abbott y otros han descubierto, en mi laboratorio, que este sistema descendente tiene mayor impacto sobre el sistema lateral que sobre el medial, lo que sugiere que gran parte de la acción de la morfina contra el dolor repentino o fásico es transmitida por los tractos neurales descendentes.

La actividad analgésica de la morfina no está, por supuesto, limitada a las astas dorsales de la médula y al cerebro medio. Se considera, por ejemplo, que la morfina interviene también en el sistema límbico, el cual, como se sabe, desempeña un papel importante en las sensaciones de dolor y de placer. Tal actividad podría muy bien amortiguar las sensaciones dolorosas producidas por el sistema medial (dolor tónico), que envía numerosísimos impulsos al sistema límbico.

Un estudio reciente, realizado por mi alumno S. Robin Cohen y por mí mismo, presta un apoyo adicional a la idea de que el influjo de la morfina sobre el sistema medial deriva, en parte, de su actividad por encima del cerebro medio. Inyectamos morfina en la habénula, pequeña porción del cerebro (situada justo detrás del tálamo) estrechamente asociada al sistema límbico y a la parte del sistema medial que hay en el cerebro medio. Las inyecciones producían analgesia en el test de la formalina, pero no en el del estirón de pata (similar al del tirón de cola), lo que sugiere que la morfina actúa en la habénula y que, al hacerlo, inhibe el sistema medial, pero no el lateral. Este y otros hallazgos indican que habrá que investigar más las áreas situadas por encima del cerebro medio, si se quiere obtener una cabal comprensión de cómo la morfina mitiga el dolor tónico, persistente, sin que las repetidas dosis de la droga induzcan tolerancia.

En vista de la complejidad de los mecanismos neurales del dolor, no ha de sorprender que la capacidad analgésica de la morfina resulte muy variable en cada persona. Una importante conclusión deducible de los estudios que se han hecho sobre esa variedad es la de que la necesidad de una dosis elevada no es, forzosamente, signo de adicción.

En uno de esos estudios, con pacientes de cáncer, Robert Kaiko, ahora en la Purdue Frederick Company en Norwalk, y sus colegas del Instituto Oncológico Memorial Sloan-Kettering, comprobaron que, para lograr un determinado nivel de analgesia, los pacientes de más edad necesitaban menos morfina que los más jóvenes, y los negros necesitaban menos también que los blancos. Asimismo, los pacientes que tenían un dolor difuso y menos vivo demandaban menos morfina que los que lo sentían muy agudo, y aquellos a quienes les dolía el estómago necesitaban menos morfina que los que tenían dolores en el pecho o en los brazos.

Los factores genéticos podrían influir también en la respuesta de cada individuo al poder analgésico de los narcóticos, como Anthony L. Vaccarino (discípulo mío), R. Andrew y R. Tasker, actualmente en la Universidad de la Isla del Príncipe Eduardo, y yo observamos hace poco al examinar los efectos de la morfina y de su antagónica la naloxona en una estirpe de ratones especialmente criada para estudios de la función inmune. Inesperadamente nos encontramos con que la sustancia "antagónica" aumentaba en realidad la analgesia de la morfina y producía, de por sí, analgesia en ratas sometidas al test de la formalina. Tan sorprendentes hallazgos, que hasta la fecha sólo han sido documentados para esta cepa de ratones, son claramente el resultado de una anomalía genética.

El descubrimiento de un influjo genético sobre los efectos de la morfina permite concebir como posible que la susceptibilidad de algunos individuos a la adicción tenga también un componente genético. Las pruebas aportadas por otros equipos concuerdan con esta idea, aunque en pocos trabajos se aborda el problema directamente.

No se puede identificar a los pacientes que acaso estén genéticamente predispuestos a la adicción morfínica, pero he de insistir una vez más en que el historial psicológico de una persona puede ser indicativo de ese riesgo. Más del 50 por ciento de los consumidores de narcóticos han tenido ataques de depresión grave, y el 87 por ciento presentan un historial de trastornos psiquiátricos.

La sociedad no distingue entre los adictos que padecen desajustes emocionales y los que, con la psique sana, sólo pretenden aliviar su sufrimiento; confusión que ha afectado a todos los sectores de la población. El ejemplo más angustioso quizá sea el del dolor innecesario en los niños.

Muchos médicos y enfermeras escatiman la aplicación de analgésicos a los pacientes menores de edad, no sólo por miedo a la adicción, sino también por la errónea creencia de que los niños y los adolescentes no sienten los dolores con tanta intensidad como los adultos. En un estudio clásico, Joann M. Eland y Jane E. Anderson, de la Universidad de Iowa, averiguaron, en 1977, que, a más de la mitad de los niños de cuatro a ocho años de edad que habían su-

frido operaciones de cirugía mayor –entre ellas la amputación de un miembro, la excisión de un tumor o intervenciones en el corazón–, no se les daba medicación alguna con la que se pudiesen aliviar sus dolores postoperatorios; y el resto de los niños recibían dosis insuficientes. Comparando los casos de 18 de los niños con los de adultos que habían sido objeto de operaciones similares, se vio que a todo el grupo de los niños se le había dado un total de 24 dosis de drogas analgésicas, mientras que a los adultos se les habían concedido en total 671 dosis.

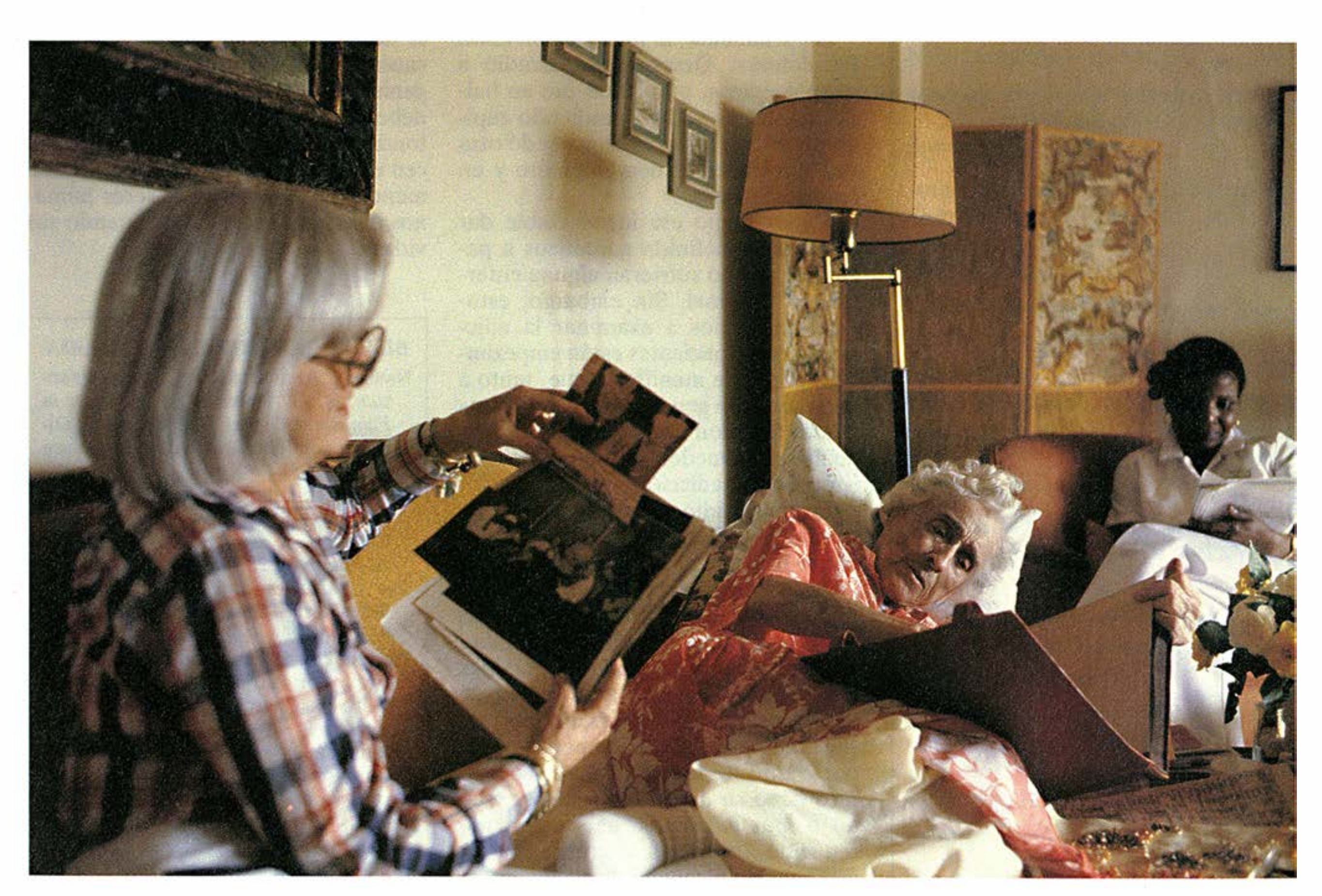
Pero también los mayores pagan las consecuencias de la ignorancia. En un estudio del dolor postoperatorio, mis colegas y yo comprobamos que en las salas de enfermos operados hay básicamente dos grupos de sujetos: el de los jóvenes y de mediana edad, que se restablecen pronto, y el de los de edad más avanzada, cuyos dolores intensos perduran y se prolongan bastante más allá del período normal de recuperación, que suele ser de tres o cuatro días. A pesar de la persistencia y del alto nivel de dolor que estos pacientes sufren (probablemente a causa de complicaciones postoperatorias) y a pesar del largo período de recuperación, no reciben dosis superiores de analgésicos ni una medicación diaria más frecuente. Alrededor del 30 por ciento de los pacientes que se encuentran en un momento dado en una sala de operados pertenecen al grupo de avanzada edad; representan, por tanto, una cantidad importante de personas que están sufriendo innecesariamente altos niveles de dolor.

Es cosa sabida que los dolores que producen las quemaduras son de los más angustiosos e insoportables, y, sin embargo, también éstos tienden a ser escasamente controlados. Manon Choinière, del Centro de Quemados del Hôtel Dieu de Montreal, y yo descubrimos que hasta en los mejores hospitales de este tipo –en los que cuentan con médicos, enfermeras, fisioterapeutas y demás personal asistencial sumamente capacitado, atento y compasivo– son altos los niveles de dolor. En nuestro estudio de una serie de 30 pacientes que habían tenido que ser sometidos a desbridación y fueron objeto de fisioterapia (ejercicios para impedir el anquilosamiento de las articulaciones) clasificamos la intensidad del dolor basándonos en el cuestionario sobre el dolor que compusimos Torgerson y yo. Supimos así que, durante las dos primeras semanas del tratamiento, el 23 por ciento de los pacientes tuvieron dolor agudo ("horrible") y el 30 por ciento dolor agudísimo ("insoportable").

Incluso en reposo, el 13 por ciento de los enfermos sentía dolor agudo, y, agudísimo, otro 20 por ciento. Estos datos, por cierto, los obtuvimos de pacientes que fueron ya medicados según las recomendaciones de los manuales clínicos al uso (o sea, ateniéndose, al prescribir las drogas, a la limitación de las siglas "PRN").

En el caso de muchos pacientes a los que se hospitaliza por operación, quemaduras o con un cáncer terminal, está claro lo que se deberá hacer: potenciar al máximo la eficiencia de la narcoterapia mediante una dosificación con enfoque preventivo del dolor. ¿Qué hay que hacer, empero, con quienes sufren un extenuante dolor crónico, pero no padecen ninguna enfermedad fatal? Estas personas han sido tradicionalmente excluidas de toda narcoterapia a largo plazo, por el temor, una vez más, a que se generara adicción.

Citaré el caso de un atleta de 26



**5. ENFERMA DE CÁNCER,** con una asistente sanitaria que la acompaña y atiende y la enfermera que la visita para verificar el seguimiento del programa hospitalario. En tales programas se procura paliar o mitigar el dolor, las náuseas y otras molestias que suelen padecer los enfermos terminales. La mayoría de los hospitales van ya adoptando el plan preventivo del dolor y se deja a los pacientes pasar en sus hogares sus últimos días.

años, que quedó inválido al producirse una grave lesión en la médula espinal y padecía terribles dolores en la espalda y en las piernas. Su sufrimiento le incapacitaba para todo tipo de trabajo, con lo que llegó a ser una carga para sí mismo, para su familia y para la sociedad, que costeaba sus gastos médicos. El doctor que le atendía descubrió que pequeñas dosis de morfina tomadas oralmente a diario (como las que reciben los enfermos de cáncer) eliminaban el dolor. Con la ayuda de tal medicación, el joven pudo volver a trabajar e hizo planes para casarse con la chica que venía siendo su amor desde la infancia y que aceptaba su invalidez.

Pero he aquí que un buen día el médico fue acusado, por el colegio de la región, de prescribir narcóticos para fines no aprobados por dicha institución y de que estaba convirtiendo a su paciente en un adicto a la morfina. Temeroso de perder su licencia médica, el doctor dejó de recetar la droga. (Donde la administración de morfina está permitida por la ley, los médicos pueden técnicamente prescribirla a voluntad, pero se encuentran coartados de hecho por los reglamentos de las sociedades médicas, que controlan la concesión de licencias para el ejercicio de la medicina.)

Naturalmente, al joven le volvieron los dolores. Desesperado, acudió a otros médicos, pero todo fue en balde. Entonces se fue hundiendo rápidamente en la depresión y quedó otra vez sumido en el desvalimiento y en la desesperanza.

Hace tiempo era inconcebible dar de forma indefinida narcóticos a pacientes que no sufrieran alguna enfermedad terminal. Sin embargo, estudios orientados a examinar la adicción en esos pacientes están empezando a poner de manifiesto que, tanto a ellos como en general a los candidatos a la terapia con narcóticos, las drogas analgésicas pueden ser de gran ayuda sin producir adicción.

En un estudio reciente, Russell K. Portenoy y Kathleen M. Foley, del Sloan-Kettering, mantuvieron a 38 pacientes no cancerosos en un tratamiento con narcóticos para evitarles sus crónicos dolores agudos; la mitad de estos pacientes recibieron opioides durante cuatro o más años, y seis de ellos fueron tratados durante más de siete años. Aproximadamente el 60 por ciento de los 38 pacientes declararon que su dolor había desaparecido o, por lo menos, había quedado reducido a un nivel tolerable. La terapia resultó problemática tan sólo en dos pacientes, que tenían ambos un historial de consumo de drogas.

Con cauto optimismo, Portenoy y Foley sugieren que la morfina pudiera ser un tratamiento razonable para evitarles el dolor crónico en muchos pacientes que no tienen enfermedades terminales. Apuntan los problemas que acaso conlleve el mantenimiento prolongado de la narcoterapia y proporcionan directrices y prudentes consejos para el control de los pacientes. Este tipo de estudios suponen para el progreso de la medicina tanto como para la aeronáutica el haber superado la barrera del sonido. Representan, en efecto, un gran paso hacia el examen razonado e imparcial de la eficacia de los narcóticos en pacientes que rara vez han podido contar con esa terapia.

Entre los críticos de la narcoterapia prolongada para esos pacientes se encuentran los médicos y otras personas que temen que a la gente se le recete sin más una droga y no se la haga partícipe de las ventajas de un planteamiento multidisciplinar para el tratamiento del dolor. Pero lo cierto es que los dos enfoques son compatibles; en realidad, se complementan recíprocamente.

Para el futuro, hacen falta muchos más estudios bien controlados que suministren datos sobre los efectos a largo plazo de los narcóticos en el dolor crónico de los enfermos no cancerosos. Al mismo tiempo, los organismos médicos y gubernamentales deberán dar las autorizaciones y los fondos necesarios para que se realicen esos estudios. El objetivo es nada menos que librar a unos seres humanos del dolor que está arruinando sus vidas.

BIBLIOGRAFIA COMPLEMENTARIA

NARCOTIC ANALGESICS IN CLINICAL PRACTICE. R. G. Twycross en *Advances in Pain Research and Therapy*, vol. 5. Dirigido por John J. Bonica et al. Raven Press, 1983.

CHRONIC USE OF OPIOID ANALGESICS IN NON-MALIGNANT PAIN: REPORT OF 38 CASES. R. K. Portenoy y K. M. Foley en *Pain*, vol. 25, págs. 171-186; 1986.

THE CHALLENGE OF PAIN. Edición revisada. Ronald Melzack y Patrick Wall. Penguin USA, 1989.

TEXTBOOK OF PAIN. Segunda edición. Dirigido por Patrick D. Wall y Ronald Melzack. Churchill Livingstone, Inc., 1989.

INFLUENCE OF THE PAIN AND SYMPTOM CONTROL TEAM (PSCT) ON THE PATTERNS OF TREATMENT OF PAIN AND OTHER SYMPTOMS IN A CANCER CENTER. Eduardo Bruera, Carleen Brenneis, Mary Michaud y R. Neil MacDonald en *Journal of Pain and Symptom Management*, vol. 4, n.° 3, págs. 112-116; 1989.

# Sol cambiante

*La constancia de su calor y brillo es mera ilusión; la emisión solar de radiación y partículas varía. La observación sistemática está comenzando a revelar las causas de esos cambios y sus efectos sobre la Tierra*

Peter V. Foukal

A quien esté tumbado en la playa o de paseo en pleno día, el Sol abrasador le parecerá constante e invariable. La verdad es que se trata de una estrella variable. Conocemos bien el "ciclo de manchas solares", que dura 11 años y se halla ahora, así se cree, cerca de su máximo; ciclo que sólo es un aspecto de una complicada fluctuación magnética de 22 años, a lo largo de la cual varía la emisión solar de luz visible y ultravioleta, rayos X y partículas cargadas. Esas fluctuaciones pueden calentar y dilatar la atmósfera superior de la Tierra, producir auroras, inutilizar líneas de alta tensión, alterar la capa de ozono del planeta y, tal vez, influir sobre su clima. Ni siquiera esta variación cíclica se puede considerar constante, porque el Sol ha mostrado diferentes pautas de comportamiento en épocas recientes, el siglo XVII, y hay buenas razones para esperar que su comportamiento cambie de nuevo.

Esta posibilidad no limita su interés al aspecto académico, porque cualquier cambio importante en la luminosidad solar, o incluso en su nivel de actividad, podría afectar a la habitabilidad de nuestro planeta. Los debates sobre posibles cambios del entorno global tienden a centrarse en el efecto de la acción humana: implicaciones climáticas de la acumulación de gases de invernadero y destrucción de ozono por los halocarburos. La comprensión y cuantificación de esos efectos exige considerar otras causas de los cambios que se producen en el medio ambiente, en particular las variaciones a largo plazo de la emisión solar de luz y partículas cargadas. Los investigadores se afanan por determinar con mayor precisión las relaciones entre las condiciones reinantes en el Sol y la Tierra, y predecir —o establecer si es posible predecir— el curso futuro de las variaciones solares.

PETER V. FOUKAL es físico solar y presidente de Cambridge Research and Instrumentation, Inc., en Massachusetts. Su investigación en curso se centra en observaciones del Sol en el infrarrojo en el Observatorio de Kitt Peak, el desarrollo de un instrumento para medir los campos eléctricos del plasma solar en el Observatorio de Sacramento Peak y el estudio de la variación de la luminosidad solar. Foukal, que comenzó su formación en la Universidad McGill de Montreal, se doctoró por la de Manchester. Entre 1969 y 1979 fue investigador y profesor del Instituto de Tecnología de California y después en la Universidad de Harvard. Acaba de publicar *Solar Astrophysics*, un manual para estudiantes.

La primera indicación de la presumible viabilidad de tales predicciones apareció en 1843, cuando Heinrich G. Schwabe, farmacéutico alemán y observador solar aficionado, anunció que el número de manchas oscuras visibles en el disco solar parecía variar siguiendo un ciclo regular de unos 10 años. La indicación de Schwabe acerca de la existencia de un ciclo de manchas solares llegó a conocimiento de J. Rudolf Wolf, director, desde 1855, del recién establecido Observatorio de Zurich. En esa ciudad suiza, Wolf siguió el curso diario del número de manchas solares basándose en informes de una red internacional de observadores; también compiló una historia del número de manchas solares basada en archivos de los últimos 150 años. Halló un período medio corregido de unos 11,1 años para el ciclo de manchas solares, aunque período y amplitud variaban notablemente de un ciclo a otro.

La gráfica del ciclo de manchas solares desde 1610 hasta el presente revela que el número de manchas solares ha oscilado sin interrupción desde aproximadamente el año 1715 [*véase la figura 2*]. Durante los trece ciclos de datos fidedignos recogidos desde 1848, la longitud del ciclo ha variado entre los 10 y 12 años. La amplitud del ciclo se ha mostrado todavía menos regular; osciló desde un número medio anual de manchas de unas 45 en 1804 y 1818 hasta un máximo de aproximadamente 190 en 1957. El ciclo en que nos hallamos hoy puede producir el más alto número de manchas solares y el mayor nivel total de actividad nunca registrados, a juzgar por su actual comportamiento.

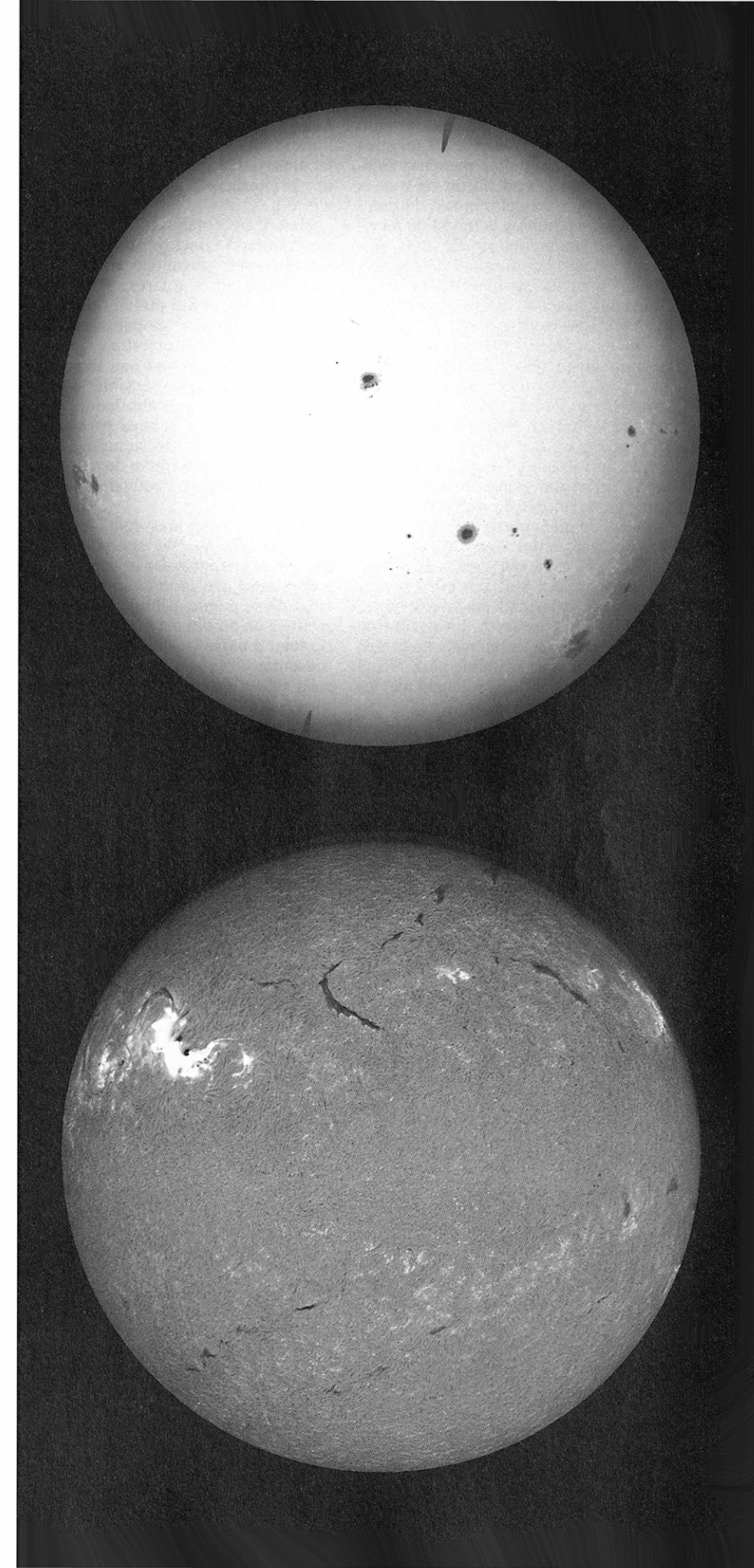
Las variaciones regulares del ciclo de manchas solares experimentan una notoria ausencia entre los años 1645 y 1715; se observaron, en ese intervalo, muy pocas manchas. A ese período de actividad solar reducida se le denomina mínimo de Maunder, en honor del astrónomo solar británico E. Walter Maunder, que se esforzó en llamar la atención sobre el mismo a fines del siglo XIX y comienzos del XX. Los astrónomos prefirieron, en un comienzo, ignorar los descubrimientos de Maunder o atribuirlos a los imperfectos telescopios y las deficientes técnicas de observación de la época.

Recientemente, sin embargo, John A. Eddy, astrónomo solar de la Corporación Universitaria de Investigación Atmosférica de Boulder, en el estado de Colorado, ha recogido pruebas sólidas de que la escasez de manchas solares documentada por Maunder constituyó un fenómeno real y bastante notable del comportamiento del Sol [véase "La desaparición de las manchas solares", por John A. Eddy; INVESTIGACIÓN Y CIENCIA, julio de 1977]. El mínimo de Maunder tuvo lugar en lo más duro de un período de tiempo insólitamente frío, la "Pequeña Edad de Hielo",

que se extendió, más o menos, del siglo XVI al XVIII; la posible relación entre ambos acontecimientos, aunque fascinante, pertenece aún al dominio de la especulación. El mínimo de Maunder y los máximos moderados de los ciclos solares al comienzo del siglo XIX podrían explicar por qué pasaron dos siglos entre las primeras observaciones europeas de manchas solares —realizadas alrededor de 1610, poco después de que Galileo y otros científicos inventaran el telescopio— y el descubrimiento de Schwabe del ciclo de 11 años.

Se sabe ahora que la variación del número de manchas cada 11 años constituye sólo el aspecto más visible de una profunda oscilación del campo magnético solar que afecta a muchos otros aspectos de la superficie y atmósfera del astro y posiblemente también a su interior. George Ellery Hale y sus colaboradores, del Observatorio del monte Wilson, en California, hallaron la primera indicación de la oscilación magnética solar en sus mediciones de los espectros de las manchas solares. Descubrieron que ciertas líneas de absorción en los espectros se ensanchaban y polarizaban de manera muy semejante a como lo hacen las líneas espectrales de gases magnetizados en el laboratorio, que habían sido estudiadas por el físico holandés Pieter Zeeman. Analizando este "efecto Zeeman", determinaron que la intensidad de los campos magnéticos en torno a las manchas se cifraba entre 2000 y 3000 gauss, miles de veces más intensos que el campo magnético terrestre. Mostraron también que la mayoría de las manchas se producían en parejas, a la manera de gigantescos dipolos magnéticos (una barra imanada es un ejemplo de dipolo), orientadas por lo general casi paralelamente al ecuador solar.

En 1912, Hale anunció que la polaridad magnética de esas parejas de manchas había cambiado de signo en las primeras manchas del nuevo ciclo que comenzó aquel año. Hacia 1924



**1. DOS VISTAS DEL SOL que revelan complejos procesos bajo su moteada apariencia. La fotografía tomada con luz visible (*arriba*) muestra manchas solares y zonas brillantes, llamadas fáculas, en la "superficie" del Sol o fotosfera. Las obtenidas en una longitud de onda del rojo, emitida por el hidrógeno, revelan detalles en la cromosfera, región más caliente y de mayor altitud (*abajo*). Puede verse una brillante fulguración, o erupción solar. Los filamentos oscuros, de gas frío y denso, permanecen suspendidos por fuerzas magnéticas. Potentes campos magnéticos determinan gran parte de la estructura externa.**

había recogido suficientes observaciones para anunciar que este cambio de polaridad se producía en cada mínimo de actividad y era una característica básica del ciclo de las manchas solares. Llegó a la conclusión de que el ciclo de 11 años del número de manchas correspondía a la mitad de un ciclo solar magnético de 22 años, durante el cual la polaridad de los grupos de manchas se invierte dos veces, volviendo así a su estado original.

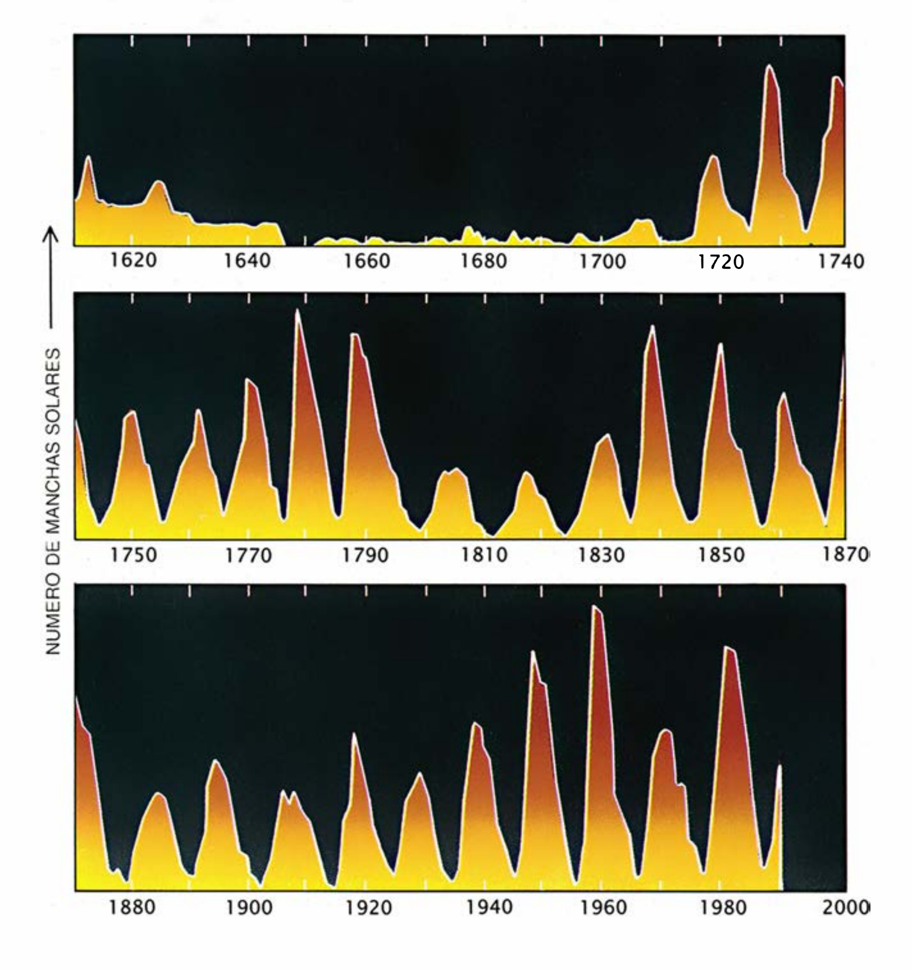
Ahora se toman diariamente mediciones mucho más sensibles del campo magnético solar mediante el magnetógrafo, desarrollado por Harold D. y Horace W. Babcock en monte Wilson en 1951. Gracias a esos magnetogramas y otras experiencias, se descubrió que el magnetismo de la superficie solar está confinado a pequeñas regiones de intenso campo magnético que cubren sólo un pequeño porcentaje del área total de la fotosfera, la capa que forma la superficie visible del Sol. Las mayores regiones de polaridad magnética única constituyen la sede de formación de manchas; otras regiones menores, cuyo tamaño va disminuyendo hasta llegar al límite de resolución de los magnetogramas más detallados (alrededor de 200 kilómetros) muestran un elevado brillo en la mayoría de las longitudes de onda de la radiación. Esas zonas superficiales de especial luminosidad reciben el nombre de fáculas (del latín, "teas"); se vieron, por primera vez, a principios del siglo XVII.

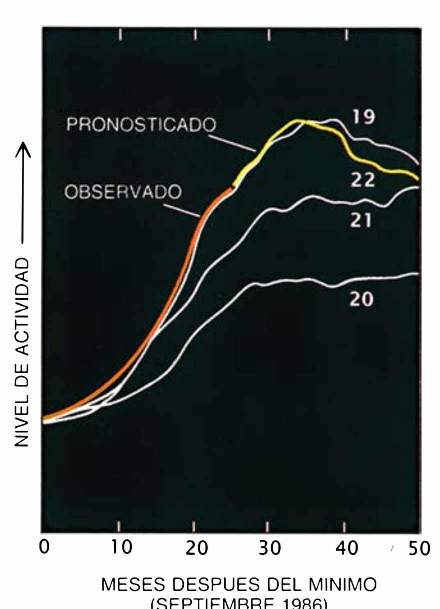
Hale y sus colaboradores trataron, sin resultado concluyente, de observar el campo magnético en la vecindad de los polos solares; pretendían determinar si el Sol poseía un campo magnético dipolar global similar al campo magnético terrestre. Durante los últimos 20 años se han llevado a cabo mediciones más sensibles que indican que los campos alrededor de los polos norte y sur solares presentan, en efecto, aunque no siempre, polaridad opuesta, y cambian de signo por la época de máxima actividad.

Esas observaciones han revelado también que la geometría del campo magnético solar es mucho más complicada que la del campo terrestre, que se puede asimilar razonablemente a un dipolo magnético. El campo solar a bajas latitudes se puede representar en forma de líneas de fuerza o tubos magnéticos arrollados en torno al Sol, siguiendo de cerca su ecuador y sumergidos bajo la superficie solar. Allí donde esas líneas de campo toroidales afloran a la superficie, forman lazadas de campo magnético que se extienden hasta las capas exteriores de la atmósfera solar, alejándose de la superficie a veces hasta millones de kilómetros en dirección a los planetas, antes de volver a conectar con el Sol. Las regiones activas, visibles en forma de manchas y fáculas, aparecen donde esas líneas cortan la fotosfera.

El mecanismo que produce el ciclo solar magnético sigue sin comprenderse bien, a pesar de la intensa investigación de que ha sido objeto durante los últimos cincuenta años. Los astrónomos suelen coincidir en que las variaciones observadas en el magnetismo solar las produce el plasma solar obligado a moverse a través de los campos magnéticos existentes. El plasma es gas altamente ionizado —gas en el que muchos electrones han sido arrancados de sus núcleos— y, por ende, buen conductor eléctrico. Los movimientos en el plasma solar inducen una corriente en el plasma y un campo magnético asociado, que refuerza, a su vez, el campo original.

A diferencia de lo que ocurre en un cuerpo sólido como la Tierra, las regiones exteriores del Sol no giran a la




**2. CICLO SOLAR. Se manifiesta en el número cambiante de manchas en la superficie visible del Sol (*izquierda*). La escasez de manchas entre aproximadamente 1645 y 1715, que definen el mínimo de Maunder, parece coincidir con una época de tiempo insólitamente frío. A tenor de las predicciones, el ciclo actual, o ciclo 22, alcanzó su máximo estos últimos meses, sobrepasado el nivel de actividad de los dos ciclos anteriores y superado el ciclo 19, el mayor (*arriba*).**

misma velocidad angular en todas las latitudes, extremo que puso ya de manifiesto el astrónomo británico Richard C. Carrington, alrededor de 1860. Las regiones ecuatoriales del Sol completan una rotación en unos 25 días, lo que significa que giran un 25 por ciento más rápidamente que los polos; entre ambas regiones, la variación es razonablemente suave. La rotación diferencial constituye, probablemente, un factor clave para impeler la dinamo que mantiene el campo magnético solar. Una línea de fuerza del campo magnético que comenzara por extenderse de manera directa a lo largo de la superficie entre ambos polos y se viera obligada a moverse con el plasma, se iría estirando, cada vez más, por la rotación ecuatorial, más rápida también. Tras unas cuantas revoluciones solares, la línea se habría arrollado de forma casi paralela al ecuador. Esta deformación de las líneas de fuerza explica quizá la geometría del campo magnético solar y la orientación este-oeste de los grupos de manchas solares [*véase la figura 4*]. El alargamiento de las líneas de fuerza del campo magnético provoca un aumento de su intensidad alcanzándose los altos valores medidos en las manchas solares.

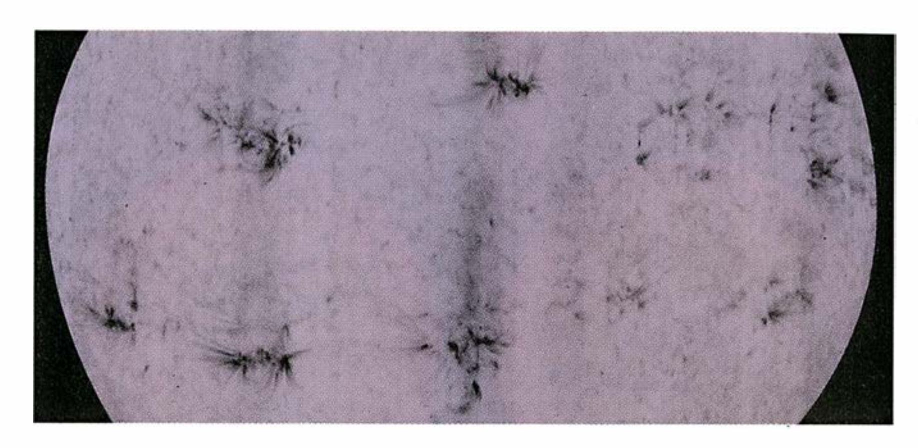
Se cree que la erupción de flujo magnético del Sol es, en parte, responsable de los cambios de polaridad de las manchas solares entre dos ciclos. Conforme los tubos magnéticos que dan lugar a las regiones activas emergen del interior del Sol, su flujo magnético se dispersa y acaba extendiéndose por la superficie del mismo. Al propio tiempo, el flujo sufre un empuje hacia fuera, hasta las capas superiores de la atmósfera solar, por los nuevos campos magnéticos que surgen desde abajo en virtud de la rotación diferencial del Sol. De ello resulta la obliteración del flujo antiguo, que algunas veces elimina la polaridad original y deja al flujo con un exceso neto de la polaridad opuesta.

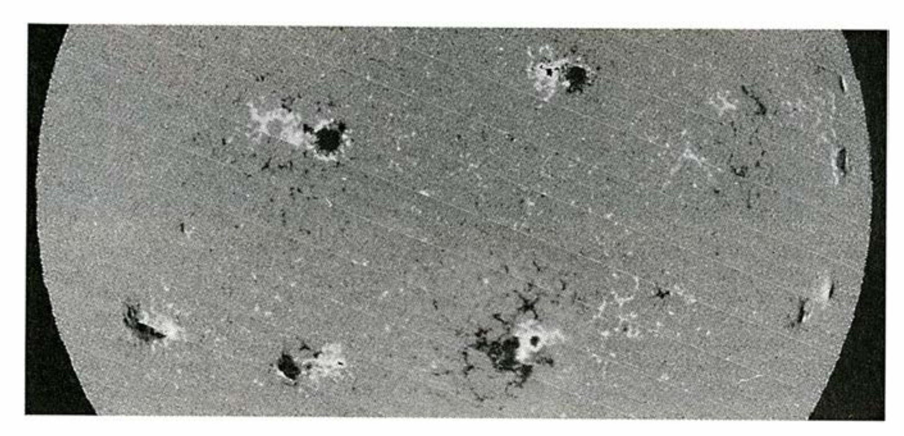
La combinación de la obliteración del "antiguo" campo magnético, por erupción y dispersión, y la generación de nuevo campo, por rotación diferencial, basta quizá para explicar el ciclo magnético solar, pero se sabe muy poco sobre el proceso mediante el cual el Sol se deshace de su antigua polaridad. Las simulaciones por ordenador de la dinámica de los movimientos del plasma solar y de sus interacciones con el campo magnético encuentran dificultades para reproducir un ciclo de 11 años en conjunción con la rotación diferencial del Sol. Una gran incógnita ha sido la naturaleza de la rotación interna del Sol, desconocida hasta hace poco. Las observaciones de las oscilaciones globales del Sol [véase "Heliosismología", por John W. Leibacher, Robert W. Noyes, Juri Toomre y Roger K. Ulrich; INVESTIGACIÓN Y CIENCIA, noviembre de 1985] están proporcionando una panorámica del interior, posibilitando el análisis del perfil de la rotación solar con la profundidad y el cálculo de su influencia sobre la dinamo solar. Descubrimientos recientes indican que el 30 por ciento exterior del interior solar gira diferencialmente, de manera semejante a la superficie. Esto sugiere que gran parte de la actividad de la dinamo solar puede residir bastante por debajo de la superficie.

Aunque los astrónomos están todavía lejos de conocer los mecanismos de la oscilación magnética del Sol, las mediciones de los efectos del ciclo magnético en emisiones solares clave constituirían en sí un importante avance. Los climatólogos quedarían satisfechos si conocieran la amplitud y escalas temporales características de las variaciones de la luminosidad solar, aun cuando la explicación astrofísica de esos cambios tuviera que esperar. Para nuestra fortuna, se han dado grandes pasos en los últimos años en esa comprensión empírica del comportamiento solar.

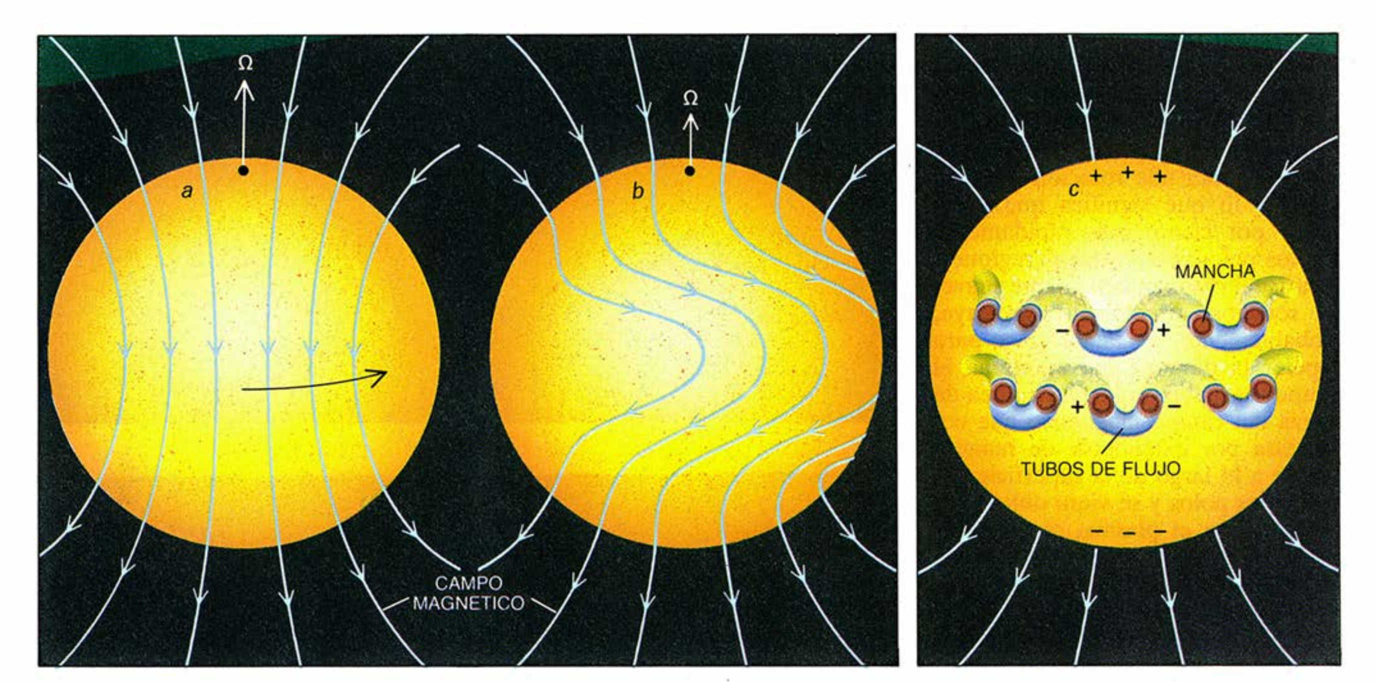
Un avance importante lo tenemos en el descubrimiento de variaciones cíclicas en la emisión total de luz solar, conocida como irradiancia solar total o —paradójicamente— "constante solar". La caprichosa atmósfera terrestre ha dificultado la medición de la constante solar, pero los instrumentos instalados a bordo de satélites revelan que llega a variar hasta en un 0,2 por ciento en la escala temporal de escasas semanas. Esta variación relativamente rápida débese al paso de manchas solares oscuras y fáculas brillantes a través del disco solar conforme el Sol efectúa su rotación más o menos mensual.

Estas fluctuaciones a corto plazo fueron identificadas —con toda nitidez, a partir de los datos obtenidos en 1980 mediante radiómetros a bordo





**3. LINEAS DE FUERZA DEL CAMPO MAGNETICO aprisionan grandes chorros de gas caliente e ionizado en la atmósfera solar, o corona. Son las líneas oscuras de las fotografías del Sol tomadas en luz ultravioleta de corta longitud de onda (*arriba*). Los chorros mayores emanan de las manchas solares. Una imagen magnética del Sol recogida al mismo tiempo, en Kitt Peak, revela intensos campos magnéticos asociados con los chorros (*abajo*). Las zonas oscuras y claras representan polaridades magnéticas opuestas; los campos más intensos aparecen en las regiones bipolares activas.**




**4. SE CREE QUE LA ROTACION DIFERENCIAL DEL SOL impulsa el ciclo magnético solar. La superficie solar arrastra con ella las líneas de fuerza del campo magnético existentes al girar el astro. Una rotación completa necesita unos 25 días en la proximidad del ecuador solar y unos 28 días en las latitudes medias; después de varias rotaciones, una línea de fuerza que siguiera inicialmente una trayectoria Norte-Sur (*a*) se ondula, se alarga horizontalmente y se intensifica (*b*). Las manchas solares aparecen donde los tubos de flujo magnético más intensos afloran a la superficie; los puntos de salida y reentrada de los tubos poseen polaridades opuestas (*c*). Las manchas solares acostumbran formarse en parejas a latitudes similares porque siguen el alargamiento horizontal del campo magnético en la proximidad de la superficie. Los pares de manchas solares constituyen enormes dipolos magnéticos, cuya orientación en el hemisferio Norte es opuesta a la que presentan en el hemisferio Sur.**

de los satélites *Misión del Máximo Solar* (*SMM*) y *Nimbus 7*– por Richard C. Willson, del Laboratorio de Propulsión a Chorro en Pasadena, y John H. Hickey, del Laboratorio Eppley en Newport, respectivamente. Más difícil fue descubrir variaciones a más largo plazo en la constante solar en el curso del ciclo de manchas, porque tales variaciones han resultado ser del orden de 0,1 por ciento; se hacía difícil descartar lentas variaciones de ese orden en la calibración del radiómetro. La prueba convincente llegó, por fin, cuando las lecturas de ambos radiómetros, que habían ido descendiendo a la par que la actividad solar desde 1980, se estabilizaron en 1986 y empezaron a aumentar en 1987 a medida que el Sol entraba en su actual ciclo y aumentaba su actividad.

Los datos de ambos satélites indican que el brillo solar decreció en aproximadamente 0,1 por ciento entre el máximo de actividad solar de 1981 y su mínimo a mediados de 1986. Para nuestra sorpresa, el Sol adquirió mayor luminosidad a medida que aumentaba el número de manchas en su superficie. El análisis de este comportamiento que hicieron Judith L. Lean, del Laboratorio Naval de Investigación, y el autor, indica que el aumento del área fotosférica cubierta por fáculas brillantes compensa con creces el aumento del área de manchas oscuras cuando crece la actividad solar.

¿Afectan esas variaciones de la constante solar al clima terrestre? Es bastante sencillo calcular el efecto de las variaciones de la irradiancia solar sobre la temperatura media global de la Tierra. Los modelos climáticos actuales sugieren que la cuantía de este efecto se halla muy por debajo de 0,1 kelvin. Esto representa una pequeña fracción del efecto de calentamiento global (varias décimas de grado) que se espera del aumento observado en la concentración del dióxido de carbono atmosférico durante los últimos decenios.

Las mediciones de irradiancia solar realizadas hasta ahora desde el espacio cubren a duras penas un ciclo de manchas solares. La información actual es insuficiente para discernir si se producen variaciones mayores de la irradiancia solar, tal vez en relación con otras variaciones más lentas de la actividad solar; por ejemplo, la que produjo el mínimo de Maunder. Se ha estimado que, durante la Pequeña Edad de Hielo, la temperatura media global llegó a descender hasta 0,5 kelvin por debajo del promedio a largo plazo, lo que bastó para desencadenar un período de apreciable avance glacial y echar a perder una serie de cosechas en Europa. Si el enfriamiento fuera el resultado de un cambio en la luminosidad solar, se habría necesitado un descenso de la irradiancia solar entre 0,2 y 0,5 por ciento actuando durante varios decenios, según cálculos basados en modelos climáticos típicos. Un programa específico de vigilancia muy precisa de la irradiancia desde el espacio exterior podría detectar variaciones lentas de la luminosidad solar y permitir que los astrónomos determinaran si una futura amortiguación de esta actividad sería susceptible de conducir a otro período dilatado de enfriamiento global.

A lo largo de los años se han multiplicado los esfuerzos por establecer relaciones entre ciclo solar y clima terrestre. El eminente astrónomo británico Sir William Herschel intuyó (¡correctamente!) que el Sol presentaba su máximo brillo durante el máximo de las manchas solares; razonó que las temperaturas resultantes, más altas, mejorarían la cosecha del trigo y harían bajar los precios. En 1801 anunció que el precio del trigo guardaba una correlación real con el ciclo de manchas solares. Pero tal correlación se esfumó y Herschel dedicó su atención a otros asuntos. Muchas otras aparentes "relaciones" han co-

nocido una vida igualmente breve, y todas ellas adolecen del mismo mal: ser relaciones estadísticas y no causales; nadie ha mostrado aún un mecanismo plausible mediante el cual tan pequeñas variaciones en la constante solar pudieran producir un efecto apreciable en la superficie terrestre.

La búsqueda, sin embargo, no se ha detenido. En 1987, Karin Labitzke, de la Universidad Libre de Berlín, informó sobre la relación más convincente hasta ahora hallada. Descubrió que la aparición de temporadas cálidas en pleno invierno en los EE.UU. y Europa Occidental mostraba una correlación notable con el ciclo solar a lo largo de los últimos 40 años, supuesto que se tenga en cuenta el cambio de dirección de los vientos estratosféricos aproximadamente cada dos años (llamado "oscilación cuasi-bienal"). La relación que encontró ha resistido repetidas pruebas estadísticas y predijo correctamente el calentamiento que produjo el invierno de 1988-1989, muy templado, en el Reino Unido y Europa Occidental. Una conexión categórica y físicamente explicable entre variabilidad solar y variabilidad climática significaría un paso de gigante en nuestro conocimiento de la relación entre el planeta y su estrella.

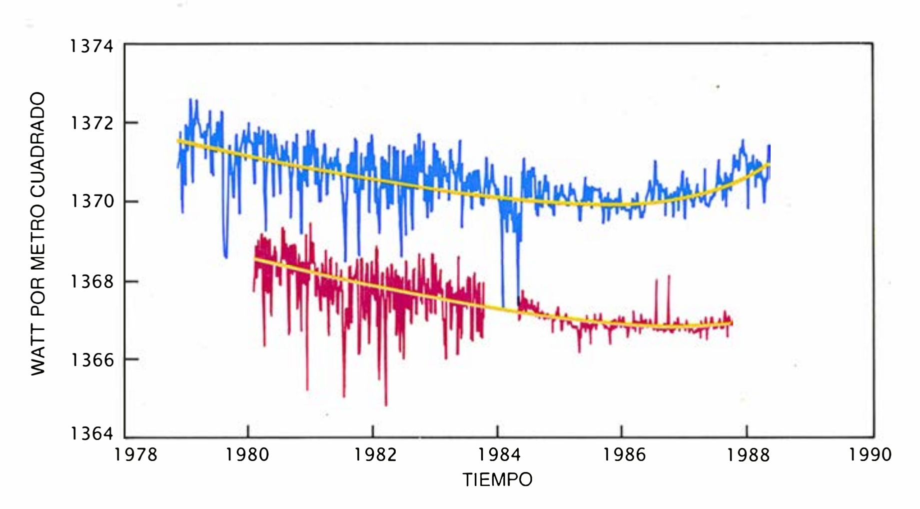
El alcance de la variabilidad solar no se circunscribe a la apariencia y brillo de la fotosfera. El ciclo magnético afecta también a capas progresivamente más altas de la atmósfera solar: cromosfera, corona y viento solar. La temperatura del plasma en esas regiones excede la de la fotosfera, aun cuando se hallen más lejos de la fuente nuclear de calor del Sol. Las pérdidas de energía por parte de esos tenues plasmas son muy bajas; pueden, por tanto, mantenerse temperaturas altísimas —de hasta varios millones de grados— con inyecciones de energía relativamente pequeñas. Esta energía deriva quizá de la disipación de ondas sonoras producidas por la agitación de la fotosfera y de corrientes eléctricas asociadas a campos magnéticos en la fotosfera o debajo de ella.

Estas capas externas y muy calientes de la atmósfera solar dan cuenta de las emisiones de rayos X, muy variables, y de la radiación ultravioleta extrema (EUV), o longitudes de onda entre unos 100 y 1000 angstrom. (Un angstrom vale $10^{-10}$ metros.) La cromosfera emite también una fracción sustancial de la radiación ultravioleta del Sol a longitudes de onda entre 1600 y 3200 angstrom; probablemente es la responsable de la mayor parte de la variabilidad de esas radiaciones. Los rayos X solares tienen una repercusión sobre la vida terrestre menor que las emisiones UV y EUV, cuyo interés obedece a su importante intervención en la atmósfera terrestre.

La causa de la variabilidad de la radiación EUV parece clara tras el análisis de las observaciones obtenidas mediante telescopios solares a bordo del *Skylab* en 1973-1974 y de satélites más recientes, como el *SMM*. Los campos magnéticos dipolares intensos y localmente cerrados en las regiones activas actúan de "jaulas" magnéticas que impiden que el plasma coronal caliente escape de la atracción gravitatoria solar. Este plasma atrapado es unas 10 veces más denso que el de las tranquilas regiones circundantes, y el plasma más denso radia mucho más intensamente en el ultravioleta extremo. Las regiones activas son, por tanto, la fuente principal de EUV solar; esas emisiones suben y bajan a la par que las variaciones de actividad en el ciclo solar.

No se ha podido medir todavía la variación de la emisión EUV en la totalidad de un ciclo de manchas solares de 11 años. Constituye un reto proyectar y construir espectrofotómetros y detectores para este dominio espectral y garantizar que, bajo la potente irradiación EUV en el espacio, su calibración permanecerá dentro de un límite que permita identificar de modo fidedigno variaciones lentas, incluso de algunas décimas de puntos de porcentaje. Las mediciones en la prominente línea de emisión del hidrógeno a 1216 angstrom (la línea Lyman alfa) indican una variación de intensidad en torno a un factor de dos.

El rápido crecimiento del presente ciclo solar actual ha producido un espectacular aumento del flujo de EUV solar, despertando preocupación en la Administración Nacional de Aeronáutica y del Espacio (NASA) por la duración de las vidas orbitales del Telescopio Espacial Hubble y del Equipamiento para Exposiciones de Larga Duración (LDEF, siglas de "Long Duration Exposure Facility"); este último es una gran plataforma orbital dotada para desarrollar 57 experimentos de ingravidez. El problema estriba en el rozamiento atmosférico: durante los períodos de alta actividad solar, el incremento del calentamiento que causa la EUV en la alta atmósfera terrestre por encima de aproximadamente 100 kilómetros puede provocar que las temperaturas aumenten en la ionosfera hasta casi tres veces los valores correspondientes a períodos de baja actividad. Las mayores temperaturas, a su vez, posibilitan que la atmósfera soporte una densidad gaseosa hasta 50 veces mayor a una altitud de 600 kilómetros, donde se supone que orbitará el telescopio espacial. La mayor densidad atmosférica producirá un mayor rozamiento y acortará la vida orbital del ingenio, que vale 1000 millones de dólares, de modo que el transbordador espacial tendría que comunicarle pronto un nuevo impulso. Ante esa situación, la NASA ha pedido a los astrónomos solares predicciones sobre la magnitud de la actividad del ciclo actual y el momento del máximo, a fin de elegir la mejor




**5. CENTELLEO del Sol, registrado por dos satélites, *Nimbus 7* (*azul*) y *Misión del Máximo Solar* (*rojo*). Las disminuciones a corto plazo en la emisión solar producen las puntas aguzadas de los datos del segundo y la mayoría de los que aparecen en los datos del *Nimbus 7*, que incorporaban también algo de ruido del instrumento. Por término medio (*línea amarilla*), el máximo brillo del Sol correspondió a la época de máxima actividad de manchas solares. Al parecer, el mayor número de fáculas brillantes en la época de máxima actividad contrarrestó los efectos de las manchas oscuras.**

fecha de lanzamiento para el telescopio espacial y preparar la próxima recuperación del LDEF.

La medición de la variación solar de 11 años resulta algo más fácil en el ultravioleta que en el EUV. Durante los últimos diez años se ha dedicado más esfuerzo a las observaciones en las longitudes de onda del ultravioleta porque afectan directamente a la capa de ozono de la atmósfera. Los datos suministrados por los satélites *Nimbus 7* y *Explorador de la Mesosfera Solar* ponen de manifiesto una variación de 27 días (causada por la rotación solar) a longitudes de onda inferiores a unos 3000 angstrom. Aunque la amplitud de la variación del ciclo solar de 11 años en el ultravioleta se muestra esquiva por culpa de las dificultades de calibración, parece encontrarse entre un máximo del 20 por ciento hacia los 1500 angstrom y sólo 1 o 2 por ciento a longitudes de onda superiores a 2500 ansgtrom.

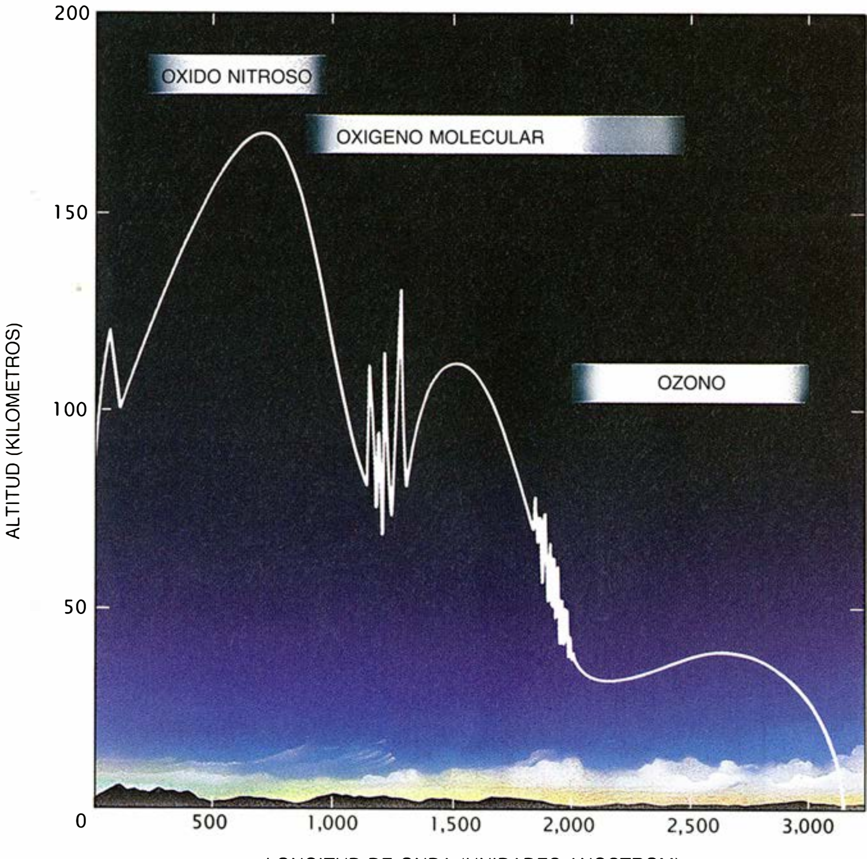
Los modelos vigentes sobre producción de ozono señalan que tales cambios en las emisiones solares en el ultravioleta podrían causar una variación del 1 al 2 por ciento en la cantidad total de ozono. Esto podría explicar, en gran parte, el decrecimiento del ozono estratosférico medido por los satélites entre 1978 y 1985, período de actividad solar principalmente decreciente. Queda claro que los efectos del ciclo solar se deben introducir en cualquier trabajo empeñado en identificar el motivo de esta disminución del ozono y de esa otra disminución, a más largo plazo, medida desde la superficie terrestre entre 1969 y 1986. La lenta disminución del ozono global es menos espectacular que el agujero recientemente identificado en la capa de ozono sobre la Antártida; ahora bien, de continuar así, será todavía más seria.

La emisión solar de partículas cargadas, que depende principalmente de las condiciones en las capas situadas por encima de la fotosfera, varía también en el curso del ciclo solar. Las más importantes de esas partículas, en virtud de su influencia sobre los sistemas terrestres, son los protones de alta energía, ocasionalmente expulsados por explosiones en la corona solar. La Tierra se ve afectada también por la expulsión generalizada de plasma coronal: el viento solar.

Los protones solares de alta energía observados en la Tierra poseen energías que varían desde unos 10 millones hasta 10.000 millones de electronvolt (en comparación, un fotón de luz visible tiene una energía de unos dos electronvolt). Los protones de energía más alta viajan a velocidades próximas a la de la luz y llegan a la Tierra unos ocho minutos después de que se hayan producido algunas de las mayores fulguraciones solares; son estas últimas unas ingentes erupciones que se producen en regiones solares activas. Durante la erupción, esas regiones brillan muchísimo más en las bandas de rayos X y de EUV. Se cree que las fulguraciones obtienen su energía de la rápida aniquilación de intensos campos magnéticos, que calienta el plasma y engendra poderosos campos eléctricos que aceleran las partículas cargadas.

Los grandes episodios de protones constituyen un motivo de preocupación para las compañías aéreas, en particular las que vuelan a lo largo de rutas polares, donde las líneas del campo magnético se curvan hacia abajo, apuntando a la superficie, y permiten que las partículas cargadas penetren hasta bajas altitudes exponiendo a los pasajeros a elevados niveles de radiación. Estos sucesos plantean una amenaza más seria para los astronautas, especialmente los que puedan volar en satélites de órbita polar. Los episodios de protones se han visto también implicados en el fallo de sistemas informáticos; el pasado mes de agosto, uno de tales episodios obligó a cerrar temporalmente la bolsa de Toronto. Durante un ciclo solar sólo ocurren algunas docenas de esas grandes fulguraciones, si bien su frecuencia es mucho mayor en la vecindad del máximo del ciclo solar que en la de su mínimo.

Las variaciones en el flujo continuo de plasma del viento solar que pasa cerca de la Tierra dan lugar a otra clase de interacciones completamente diferentes. Podemos imaginarnos ese plasma de energía baja como una suerte de "derrame" de la corona solar, que se encuentra demasiado caliente para que la atracción del campo gravitatorio solar logre retenerla en toda su integridad. El viento solar




**6. PROFUNDIDAD DE PENETRACION de la radiación solar en la atmósfera terrestre; varía con su longitud de onda. El gráfico muestra las alturas a las cuales se ha absorbido aproximadamente la mitad de una radiación. Por fortuna para la vida, el óxido nitroso de la tenue atmósfera situada a más de 50 kilómetros por encima de la superficie de la Tierra bloquea la emisión solar, altamente variable, en el ultravioleta de muy corta longitud de onda. A menores altitudes, el ozono y el oxígeno molecular absorben los rayos ultravioleta de mayores longitudes de onda, perjudiciales también para la vida. Las variaciones de la emisión ultravioleta solar afectan a la estructura de la capa de ozono.**

queda apartado de la inmediata vecindad de la Tierra por el campo magnético del planeta, que ejerce una fuerza electromagnética sobre las partículas que atraviesan sus líneas de fuerza; la región alrededor de la Tierra de la cual se halla excluido el viento solar recibe el nombre de magnetosfera. Las fulguraciones y otras erupciones magnéticas que se producen en la atmósfera solar desorganizan el viento solar y, así, se altera la presión del plasma sobre la magnetosfera.

Por lo general, las fluctuaciones resultantes del campo geomagnético sólo alcanzan un 0,1 por ciento de su intensidad, aproximadamente de un gauss. Pero las corrientes eléctricas que inducen en grandes conductores de la superficie de la Tierra, redes de líneas de alta tensión y oleoductos, por ejemplo, producen efectos devastadores. Por citar un caso, el 13 de marzo de 1989, una potente tormenta geomagnética causada por fulguraciones asociadas a una de las mayores manchas jamás observada dejó sin electricidad a toda la provincia de Quebec.

Esas potentes tormentas geomagnéticas vienen desencadenadas, en parte, por las fulguraciones que brotan en las regiones activas del Sol; tales tormentas aumentan, pues, su frecuencia con el número de manchas solares durante el ciclo magnético. El flujo más constante del viento solar parece proceder de zonas coronales que se hallan fuera de las regiones activas, donde las líneas del campo magnético solar se extienden hacia la Tierra e incluso allende la misma, abriendo así un camino a lo largo del cual las partículas cargadas viajan sin sufrir apenas perturbaciones.

En algunas regiones del Sol, la configuración de las líneas de fuerza abiertas permite que las partículas cargadas del viento solar se escapen. Fenómeno que provoca la creación de zonas empobrecidas en plasma coronal: "agujeros coronales". Siempre hay agujeros cerca de los polos solares, pero pueden desarrollarse en latitudes inferiores. Los agujeros de baja latitud producen chorros de viento solar de gran velocidad que riegan la Tierra directa y recurrentemente conforme gira el Sol. La aparición de los agujeros está ligada al ciclo solar, pero con diferente modulación de la que se observa para las manchas. Aunque sólo se dispone de datos relativos a menos de dos ciclos solares completos, los mayores agujeros a baja latitud se desarrollan, así parece, durante la fase menguante de un ciclo; significa ello que su contribución a la actividad geomagnética será mayor algunos años después de registrado el número máximo de manchas solares.

Las múltiples formas en que la variabilidad solar afecta al entorno terrestre ponen de relieve cuán útil sería poder predecir la intensidad y el calendario del próximo máximo de manchas solares. El alcance de los esfuerzos actuales en el terreno de la predicción es reducidísimo, pues se basan en reglas empíricas deducidas del comportamiento de los ciclos anteriores. A pesar de ello, esas reglas han proporcionado algunas estimaciones de interés sobre el presente ciclo —ciclo 22— y han servido de ayuda en los cálculos de la NASA para duraciones de vida orbital de satélites. Del comportamiento actual se deduce que este ciclo igualará o superará el nivel de actividad del mayor ciclo del que poseemos un registro fidedigno: el ciclo 19, que tuvo su máximo en 1957.

Para predecir la actividad solar futura resulta imprescindible conocer la ocurrida en el pasado. Varias fuentes de información nos remontan más allá incluso que la primera observación telescópica de manchas, realizada en 1610. Las observaciones a simple vista de grandes manchas datan al menos del siglo IV a. de C., aunque antes de la época de Galileo creíase que las manchas eran planetas u otros fenómenos no solares que atravesaban el disco solar. Las auroras visibles a bajas latitudes son causadas principalmente por fulguraciones solares; las observaciones de auroras han demostrado constituir un medio menos ambiguo de inferir el comportamiento del Sol en el pasado.





**7. TELESCOPIOS SOLARES y su progresiva complejidad. En el decenio de 1850, J. Rudolf Wolf realizó las primeras mediciones sistemáticas del ciclo de manchas solares usando un pequeño telescopio refractor que está todavía funcionando en el Observatorio de Zurich (*izquierda*). En los años ochenta de nuestro siglo, el satélite *Misión del Máximo Solar* (*derecha*) llevó a cabo finas mediciones de la estructura de la atmósfera solar y de sus radiaciones variables. El satélite se estrelló contra la Tierra el 2 de diciembre de 1989; su órbita fue descendiendo a causa de la dilatación de la atmósfera superior de la Tierra causada por los altos niveles de actividad solar que el satélite iba registrando.**

Existe un registro notablemente largo de la actividad solar encerrado en las concentraciones históricas del carbono 14, isótopo radiactivo del carbono 12. La producción de carbono 14 en la atmósfera terrestre está determinada por el flujo de partículas de alta energía: los rayos cósmicos galácticos, producidos en procesos de alta energía fuera del sistema solar. La intensidad y geometría de los campos magnéticos transportados desde el Sol por el viento solar cuando la actividad de éste es grande reduce la capacidad de los rayos para penetrar en el sistema solar. En el curso de la fotosíntesis, las plantas absorben C-14, junto con los otros isótopos del carbono, y lo incorporan en su estructura. Los niveles de actividad solar a lo largo del último par de milenios pueden estimarse estudiando la concentración relativa de C-14 en los anillos de árboles viejos, aunque vivos todavía; las edades de los anillos se calculan contando simplemente, desde el presente, hacia atrás.

Los resultados de antiguas observaciones visuales de manchas, las auroras que se vieron y los datos del C-14 fueron compilados en 1976 en el estudio pionero de Eddy. Este determinó, más allá de toda duda razonable, que el mínimo de Maunder coincidió con espectaculares reducciones en el nivel de actividad solar, según indicaba la escasez de auroras y los altos niveles de C-14. Eddy y otros han demostrado, además, que tales episodios de actividad solar anormalmente baja durante muchos decenios

constituyen un aspecto bastante común del comportamiento solar: otro episodio, el mínimo de Spörer, tuvo lugar aproximadamente entre 1450 y 1550. Asimismo, un dilatado período de intensa actividad solar que ocurrió hacia los años comprendidos entre 1100 y 1250 coincidió con el tiempo relativamente cálido que posibilitó, se presume, las emigraciones de los vikingos hasta Groenlandia y el Nuevo Mundo. A tenor del registro histórico, cabría razonablemente esperar, para el siglo que viene, otro período de calma en la actividad solar.

Una fuente adicional de información sobre el comportamiento "normal" del ciclo solar y la variabilidad del astro en el pasado lejano se nos ofrece en las observaciones de otras estrellas de masa similar a la del Sol. En un estudio pionero publicado en 1976, Olin C. Wilson, del Observatorio del monte Wilson, mostró que la emisiones cromosféricas observables, en luz visible, en muchas de tales estrellas presentan variaciones cíclicas con períodos semejantes a los del ciclo solar. Constituía el primer indicio firme de que los ciclos de actividad magnética eran un rasgo común de las estrellas parecidas al Sol en masa y edad (y, por tanto, en tamaño y temperatura). Las mediciones de curvas de luz estelar aportan, por su lado, pruebas provisionales de rotación diferencial en tales estrellas.

Los estudios relativos a estrellas mucho más jóvenes que el Sol muestran cuán intensas y variables pudieron haber sido la luz ultravioleta del astro y la emisión total hace miles de millones de años, cuando la vida apareció sobre la Tierra. De acuerdo con esa línea de investigación, el nivel de actividad fotosférica y cromosférica en una estrella, y de su emisión coronal de rayos X, guardan una correlación estrecha con su velocidad de rotación. Las estrellas más jóvenes, por lo general, giran más deprisa que las antiguas; por ende, su intensidad en rayos X y luz UV es mayor a la vez que más variable; su luminosidad total tiende también a variar más. Un descubrimiento, reciente e inesperado, remite a la luminosidad de las más jóvenes, que parecen brillar menos durante períodos de gran actividad, lo que implica que, para ellas, el efecto de las "manchas estelares" predomina sobre el de las fáculas, al revés que en el caso del Sol.

¿Cabe esperar con fundamento que los astrónomos predigan el comportamiento solar? Podría haber limitaciones básicas en la posibi-

lidad de tales predicciones si los procesos que impulsan el ciclo solar no fueran lineales. Los sistemas no lineales carecen del comportamiento previsible que encontramos en los osciladores simples, de los que tenemos un ejemplo en el péndulo; una realimentación relativamente directa de un "efecto" en su "causa" puede conducir a un comportamiento errático de exasperante complejidad. Aun cuando su comportamiento esté regido por un conjunto de fuerzas bien conocido, los osciladores no lineales son tan sensibles a las condiciones iniciales que las predicciones que se adentren en el futuro se tornan inviables.

En los años cincuenta, mientras investigaba la predicción del tiempo a largo plazo, Edward N. Lorenz, del Instituto de Tecnología de Massachusetts, hizo importantes contribuciones al estudio del comportamiento no lineal. Mostró que la realimentación de varios mecanismos de la atmósfera dificulta la predicción del tiempo porque los valores actuales de la presión, temperatura y viento deben conocerse con una exactitud imposible si se desea predecir con precisión más allá de unos días por delante. El análisis de la dinámica del ciclo solar sugiere que su comportamiento bastante regular entre 1700 y hoy, así como su desaparición durante seis períodos de 11 años entre 1645 y 1715, podrían ser característicos de un oscilador no lineal.

Las investigaciones en marcha deben ayudar a establecer si el ciclo solar es previsible, al menos en principio, o si es caótico. Aun cuando fuera imprevisible, el conocimiento de las posibles relaciones entre las lentas variaciones de la actividad solar y el clima revestiría interés para desenmarañar el registro del pasado climático de la Tierra y nos prepararía para los cambios que puedan aguardarnos en los siglos venideros.

BIBLIOGRAFIA COMPLEMENTARIA


SUN, WEATHER, AND CLIMATE. J. R. Herman y R. A. Goldberg. Servicio Nacional de Información Tecnológica, NASA-SP-426; 1978.

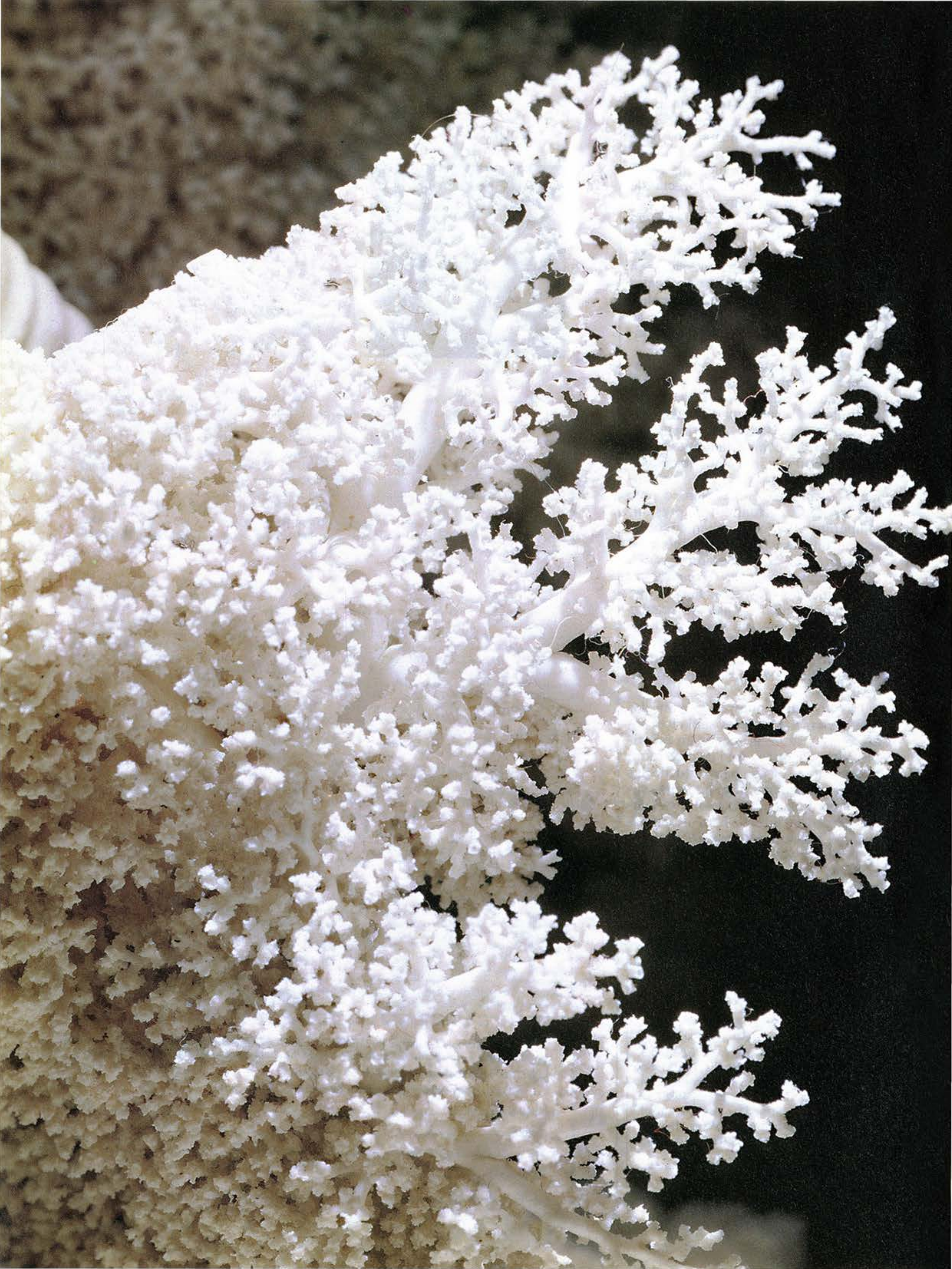
THE SUN, OUR STAR. Robert W. Noyes. Harvard University Press, 1982.

SUN AND EARTH. Herbert Friedman. Scientific American Books, Inc., 1986.

ASTROPHYSICS OF THE SUN. Harold Zirin. Cambridge University Press, 1988.

THE RESTLESS SUN. Donat G. Wentzel. Smithsonian Institution Press, 1989.

SOLAR ASTROPHYSICS. Peter Foukal. Wiley Interscience, 1990.



# Caos y fractales en la fisiología humana

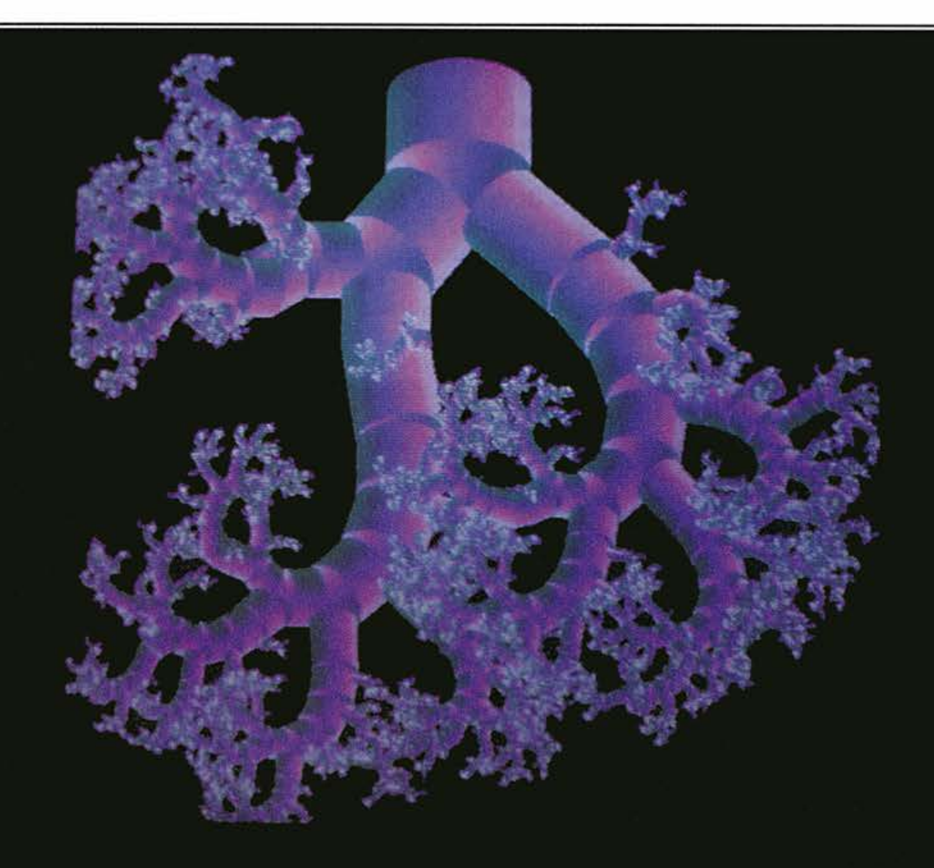
*En las funciones corporales, el caos es señal de buena salud. Un comportamiento periódico puede ser, por contra, presagio de enfermedad*

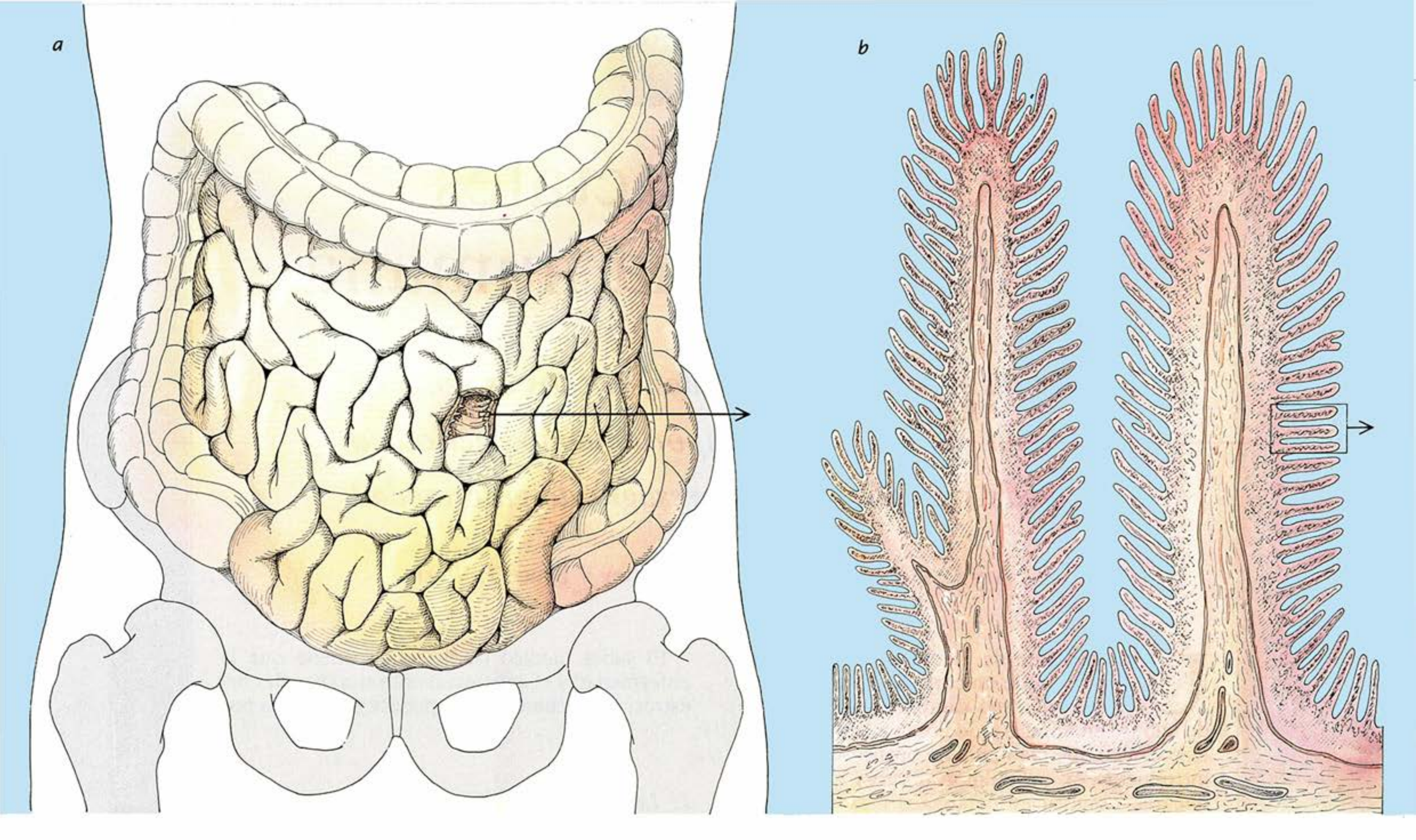
Ary L. Goldberger, David R. Rigney y Bruce J. West

Un estudiante de medicina que supervisa los ritmos de un corazón observa que, a veces, el ritmo cambia llamativamente de minuto en minuto o de hora en hora. Un médico clínico que manipula un broncoscopio para introducirlo en un pulmón observa que la tráquea se ramifica en conductos más y más pequeños. El estudiante tiene la impresión de que el intervalo entre latido y latido varía caóticamente. Quizás el clínico se percate de que la red de conductos pulmonares se asemeja a una fractal. Tan sólo recientemente han comenzado fisiólogos y médicos a cuantificar tales posibilidades de dinámica caótica y de arquitecturas fractales. Sus investigaciones están poniendo en entredicho principios médicos sostenidos desde hace largo tiempo y están revelando posibles indicios admonitorios de enfermedad.

El saber médico tradicional sostiene que la enfermedad y el envejecimiento son secuelas del estrés, consecuencias que padece un sistema por

*1. Los conductos del pulmón* (izquierda), *conformados por la evolución y el desarrollo embrionario, se parecen a fractales generados por ordenador* (abajo). *Los bronquios y bronquiolos del pulmón (en este caso, un vaciado en látex) forman un "árbol" compuesto por múltiples generaciones de ramificaciones. La ramificación a pequeña escala de los conductos se asemeja a la ramificación a escalas mayores. Cuando los fisiólogos cuantificaron sus observaciones de las pautas de ramificación, descubrieron que el árbol pulmonar posee geometría fractal.*






lo demás ordenado y cuasi-mecánico; que el estrés hace disminuir el orden, provocando respuestas erráticas o alterando los ritmos periódicos y normales del organismo. A lo largo de los cinco últimos años, nuestros colegas y nosotros hemos descubierto que el corazón y otros sistemas fisiológicos pueden comportarse de modos sumamente erráticos cuando son jóvenes y sanos, y que, contrariamente a lo que podría sugerir la intuición, el envejecimiento y la enfermedad van acompañados a veces de comportamientos de regularidad creciente.

La irregularidad y la impredictibilidad son, estamos viendo, rasgos característicos y señalados de la salud. Por otra parte, la disminución de variabilidad y la acentuación de las periodicidades son factores asociados a la enfermedad. Motivados por estas ideas, hemos buscado, lo mismo que otros fisiólogos, comportamientos periódicos que pudieran ser indicativos del desarrollo de dolencias (enfermedades cardíacas, sobre todo). Además, hemos empezado a analizar la flexibilidad y la fortaleza de las estructuras fractales y la adaptabilidad y robustez de sistemas que, en una primera impresión, exhiben conductas caóticas.

ARY L. GOLDBERGER, DAVID R. RIGNEY y BRUCE J. WEST han colaborado en estudios de dinámica no lineal en fisiología. Goldberger es profesor de medicina en la Facultad de Medicina de Harvard y codirector de los laboratorios de electrocardiografía y arritmia del Hospital Beth Israel, en Boston. Rigney es profesor de medicina en Harvard e investigador del Instituto de Tecnología de Massachusetts. West preside el departamento de física de la Universidad del Norte de Texas.

Las fractales y el caos son materias asociadas con una disciplina matemática llamada dinámica de sistemas no lineales, es decir, de sistemas que no responden a los estímulos en proporción directa. La teoría de sistemas dinámicos no lineales ayuda a comprender el fenómeno de las epidemias, la cinética de ciertas reacciones químicas y los cambios climatológicos. En ciertas circunstancias, sistemas determinísticos no lineales —los que poseen sólo unos pocos elementos simples— se comportan erráticamente, entrando en un estado llamado caos. No hemos de confundir el caos determinístico de la dinámica no lineal con el caos que el diccionario define como aleatoriedad o desorganización completa. El caos no lineal alude a una aleatoriedad de clase restringida, una aleatoriedad que, llamativamente, podríamos asociar con la geometría fractal.

Las estructuras fractales son, a menudo, vestigios o residuos producidos por los sistemas dinámicos no lineales. Cada vez que un proceso caótico ha dado forma a un ambiente (sea la costa, la atmósfera o una falla geológica) es verosímil que haya dejado en pos de sí estructuras fractales (sea el contorno del litoral, las nubes o las formaciones rocosas). No obstante lo cual, la matemática de los fractales se desarrolló con independencia de la dinámica no lineal, y ni siquiera en nuestros días se encuentran plenamente establecidas las conexiones entre una y otra disciplinas.

Un fractal, cual fue inicialmente concebido por Benoit B. Mandelbrot, del Centro de Investigación Thomas J. Watson de IBM, consta de fragmentos geométricos de orientación y tamaño variable, pero de aspecto similar. Ciertas neuronas, por ejemplo, poseen una estructura fractaliforme. Al examinar tales neuronas mediante un microscopio de limitada potencia, resultan discernibles ramas asimétricas, llamadas dendritas, conectadas a los cuerpos celulares. Bajo ampliación ligeramente mayor, se observan sobre las anteriores otras ramas menores. A mayor amplificación todavía, vemos otro nivel de detalle: ramas que emanan de ramas que parten de ramas. Aunque la ramificación de las neuronas llega a detenerse a cierto nivel, los fractales ideales poseen infinito detalle.

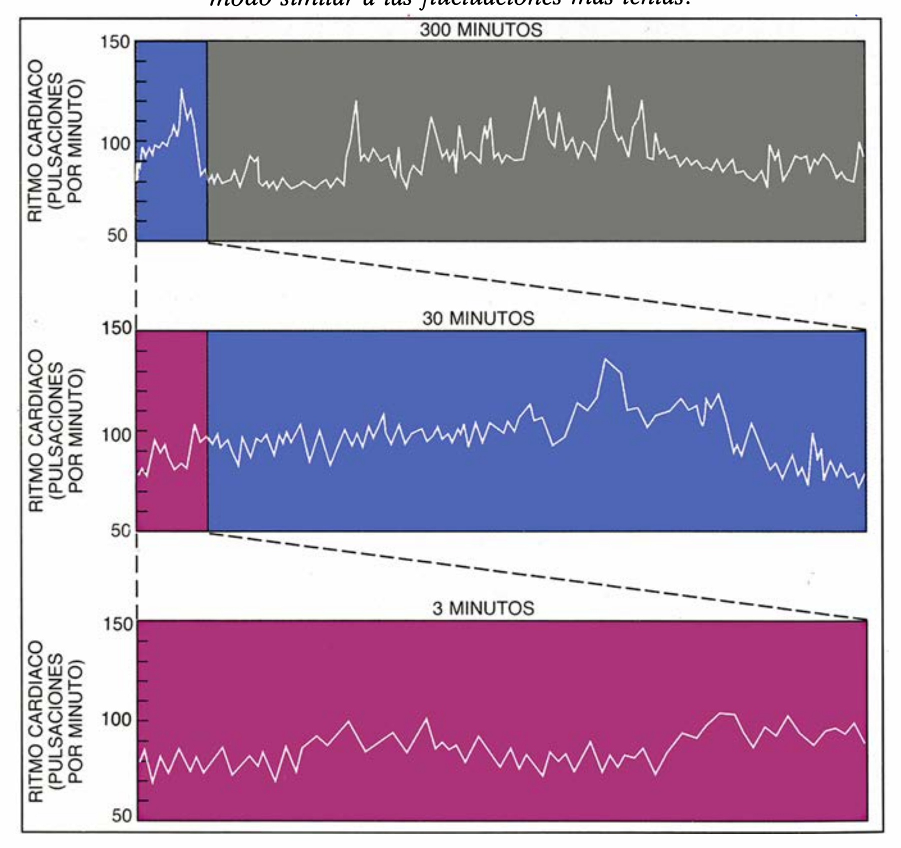
Tal vez sea más destacable aún que los detalles de un fractal a cierta es-

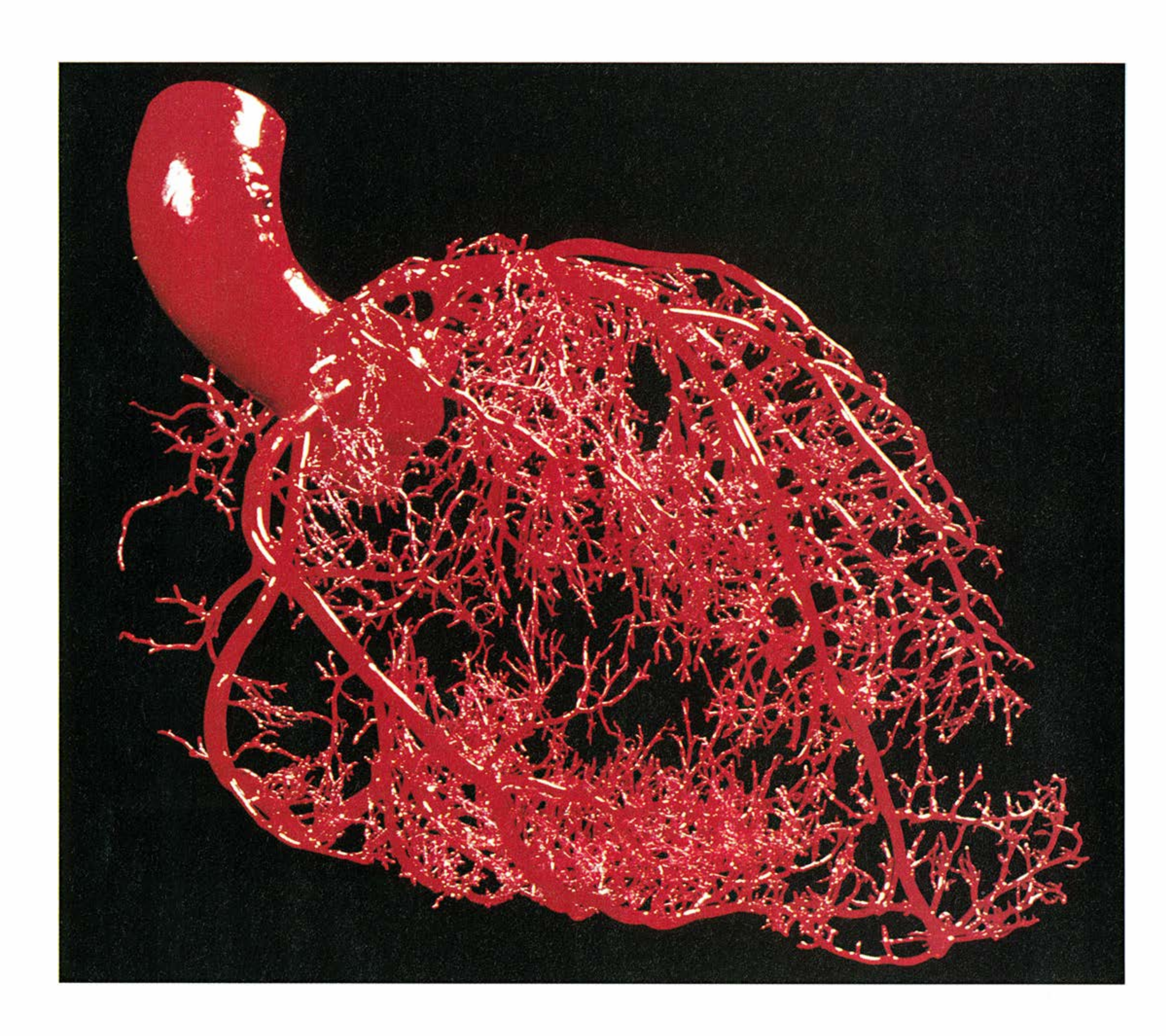
cala sean semejantes (aunque no necesariamente idénticos) a los de la estructura visible a escala mayor o menor. Si observásemos dos fotografías de las dendritas tomadas a ampliaciones diferentes (sin ningún elemento ajeno que sirviera de referencia), resultaría difícil decidir a qué ampliación se tomó cada una. Todos los fractales poseen esta propiedad interna de parecerse a sí mismos, propiedad que recibe el nombre de sibisemejanza.

Dado que un fractal está compuesto por estructuras similares de "grano" cada vez más fino, su longitud no está taxativamente definida. Cuando se intenta medir la longitud de una línea fractal con una determinada regla, algunos detalles serán siempre más finos de lo que la regla tiene posibilidad de medir. Así, pues, conforme aumenta la resolución del instrumento de medida, va creciendo también la longitud de un fractal.

La noción de longitud carece de significado en el caso de los fractales. Por ello, los matemáticos han ideado un número, al que llaman "dimensión", para cuantificar de qué modo llena el espacio un fractal. El concepto de dimensión que nos es familiar se aplica a los objetos de la geometría clásica, euclídea. Las rectas tienen dimensión uno, los círculos son de dimensión dos y las esferas lo son de tres. Pero los fractales tienen

*2. Sibisemejanza de un sistema, en virtud de la cual las características de una estructura o proceso ofrecen el mismo aspecto a diversas escalas, sean de longitud o de tiempo. Cuando se observan las estructuras del intestino delgado a diferentes ampliaciones (en los dibujos que hay sobre estas líneas), el parecido entre los detalles mayores y los más pequeños hacen pensar en sibisemejanza. Cuando se registra el ritmo cardíaco de un individuo sano durante 3, 30 y 300 minutos* (curvas al pie), *las fluctuaciones rápidas y erráticas parecen variar de modo similar a las fluctuaciones más lentas.*

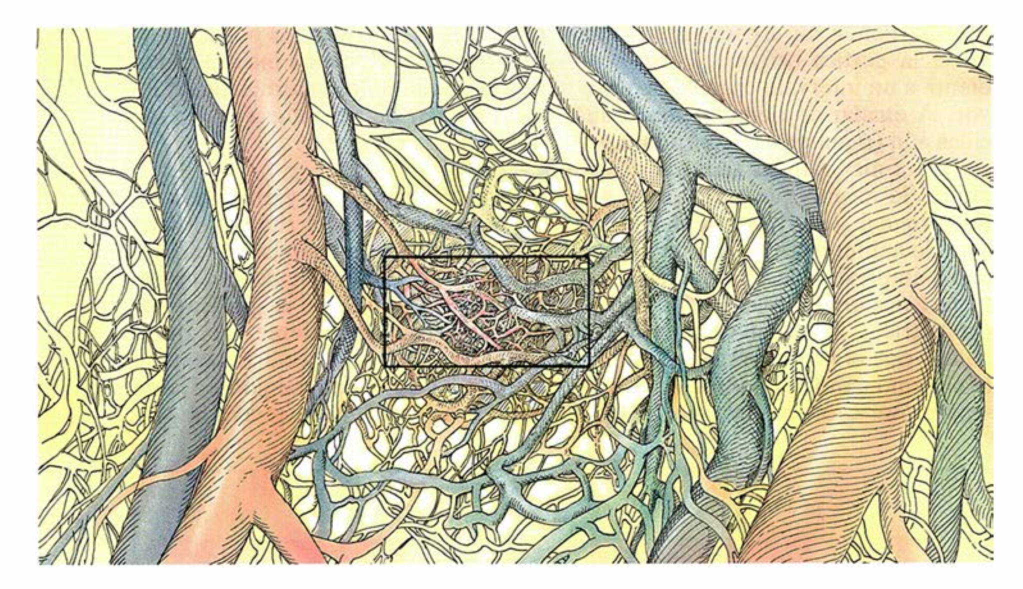


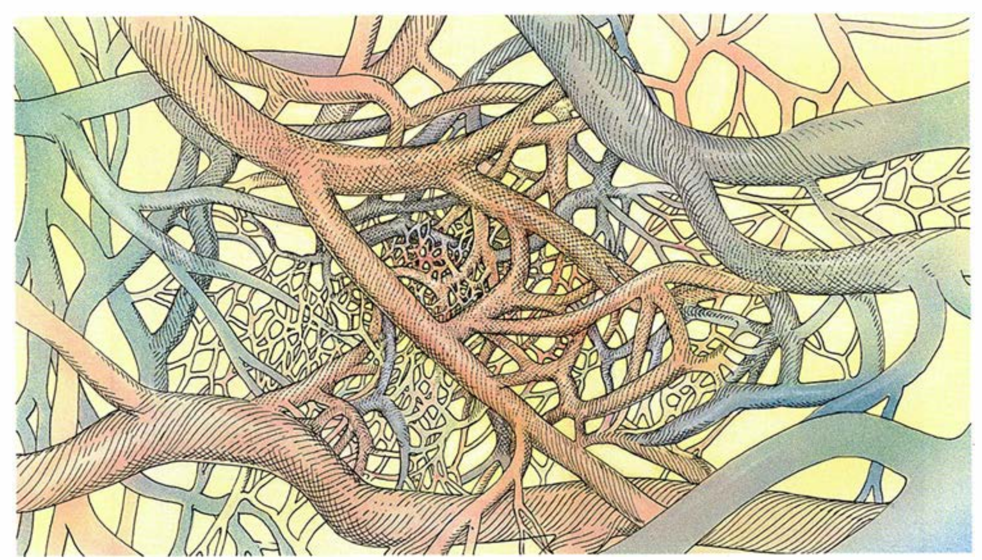


dimensión no entera (fraccionaria). Mientras que una línea euclídea lisa llena exacta y precisamente un espacio unidimensional, una línea fractal se desparrama en un espacio bidimensional. En consecuencia, una línea fractal —la de la costa, por ejemplo— tiene dimensión comprendida entre uno y dos. Análogamente, una superficie fractal —una montaña, verbigracia— tiene dimensión intermedia entre dos y tres. Cuanto mayor es la dimensión de un objeto fractal, tanto mayor es la probabilidad de que una región dada del espacio contenga una porción del objeto fractal.

Abundan en el cuerpo humano las estructuras fractaliformes, que podemos observar en las redes nerviosas, las redes de vasos sanguíneos y las de otros conductos. El sistema fractal más cuidadosamente estudiado del cuerpo humano es el sistema de tubos pulmonares encargados del transporte y la evacuación de gases. En 1962, Ewald R. Weibel y Domingo M. Gómez (y más tarde, Otto G. Raabe y colegas) efectuaron mediciones cuidadosas de las longitudes y diámetros de los conductos de este irregular sistema de vías aéreas. Recientemente, dos de los autores (West y Goldberger), en colaboración con Valmik Bhargava y Thomas R. Nelson, de la Universidad de California en San Diego, analizaron nuevamente esas mediciones, valiéndose de vaciados de pulmones humanos y de otras diversas especies de mamíferos. Descubrimos, a pesar de sutiles diferencias entre unas y otras especies, el tipo de factor de escala que correspondería a lo predicho para las dimensiones de un objeto fractal.

Muchos otros órganos prometen también ser fractales, mas sus dimensiones están todavía por cuantificar. Las estructuras fractaliformes desempeñan un papel vital en la dinámica eléctrica y mecánica del corazón sano. En primer lugar tenemos, por ejemplo, una red fractaliforme de arterias y venas coronarias que aportan sangre a los músculos cardíacos y la evacuan de ellos. Hans van Beek y James B. Bassingthwaighte, de la Universidad de Washington, recurrieron recientemente a la geometría fractal para explicar anomalías en las pautas de flujo sanguíneo que penetra en el corazón sano. La interrupción de este flujo arterial puede provocar





*3. Vasos sanguíneos del corazón. Manifiestan una ramificación fractaliforme. Los vasos mayores* (vaciado de la izquierda) *se escinden en vasos más finos* (dibujo superior), *los cuales se ramifican, a su vez, en vasos aún más finos* (dibujo inferior).

el infarto de miocardio (ataque al corazón). En segundo lugar, una envoltura fractaliforme de fibras de tejido conjuntivo —las cuerdas tendinosas— actúa a modo de traílla que apersoga las válvulas mitral y tricúspide a la musculatura cardíaca subyacente. La ruptura de estos tejidos puede producir una regurgitación muy seria de sangre desde los ventrículos a los atrios, seguida de fracaso cardíaco congestivo. Por último, la arquitectura fractal también es evidente en las pautas de ramificación de ciertos músculos del corazón, así como en el sistema His-Purkinje, que conduce impulsos eléctricos cardíacos de los atrios a los músculos cardíacos de los ventrículos.

Aunque las organizaciones anatómicas fractales anteriores sirven a funciones, en apariencia dispares, en diferentes sistemas orgánicos, salen a la luz diversos temas anatómicos y fisiológicos comunes. Las ramificaciones y repliegues fractales amplifican grandemente la superficie de las áreas de absorción (como el intestino), de distribución o recolección (por los vasos sanguíneos, los conductos biliares o los tubos bronquiales) y de procesamiento de información (por los nervios). Las estructuras fractales, a causa, en parte, de su redundancia e irregularidad, son robustas y resistentes a las lesiones. El corazón, por ejemplo, puede continuar bombeando con disfunción mecánica relativamente mínima, aun cuando su sistema His-Purkinje (que se encarga de conducir los impulsos eléctricos cardíacos) haya sufrido grandes daños.

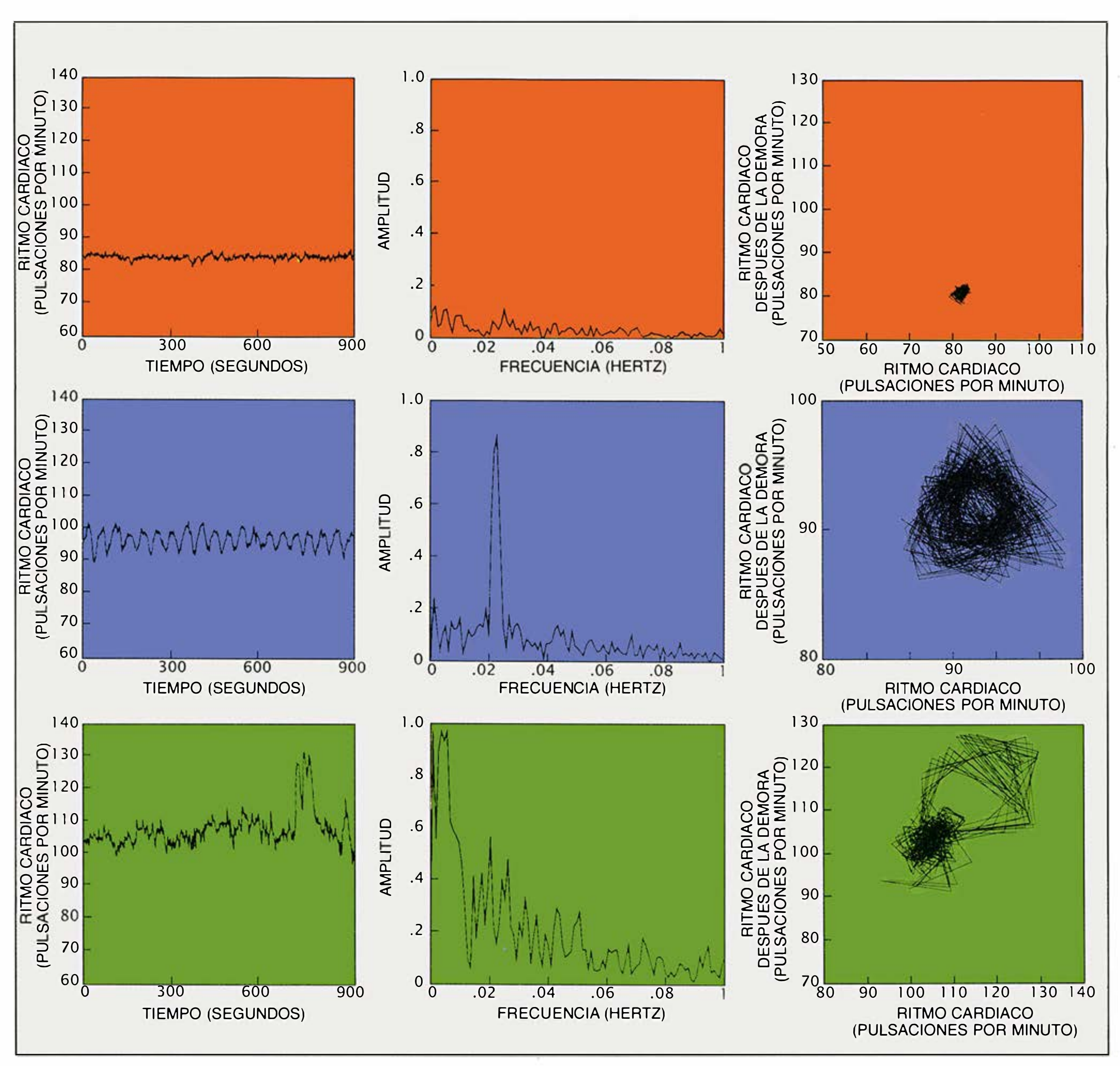
Las estructuras fractales del cuerpo humano surgen como resultado de la lenta dinámica del desarrollo y evolución embrionarios. Hemos sugerido que estos procesos —al igual que otros que engendran estructuras fractales— exhiben una suerte de caos determinístico. Investigaciones recientes en fisiología han descubierto otros ejemplos de dinámica manifiestamente caótica a escalas de tiempo más cortas y accesibles a la experimentación. A comienzos del decenio de 1980, cuando los investigadores comenzaron a aplicar las teorías de caos a los sistemas fisiológicos, esperaban que aquél se mostrase, sobre todo, en los sistemas envejecidos o enfermos. De hecho, tanto la intuición como la tradición médica les daban buenos motivos para creerlo así. Si se escuchan los latidos del corazón mediante un estetoscopio o se toma el pulso en la muñeca, el ritmo del corazón parece ser regular. En efecto, en un individuo en reposo, la intensidad del pulso y el intervalo entre latidos parecen sensiblemente constantes. Por esta razón, los cardiólogos ven en el ritmo cardíaco normal un ritmo regular senoideo.

Un análisis más minucioso revela, en cambio, que los ritmos cardíacos de los individuos sanos fluctúan considerablemente, incluso en reposo. En adultos jóvenes y sanos, el ritmo cardíaco, que es por término medio de 60 pulsaciones por minuto, puede variar hasta 20 pulsaciones por minuto cada pocos latidos. A lo largo del día, el ritmo cardíaco puede variar desde 40 hasta 180 pulsaciones por minuto.

Desde hace no menos de 50 años, los médicos han interpretado las fluctuaciones en el ritmo cardíaco recurriendo al principio de homeostasis: los sistemas fisiológicos operan normalmente en el sentido de reducir la variabilidad y de mantener una constancia en las funciones internas. De acuerdo con esta teoría, desarrollada por Walter B. Cannon, de la Facultad de Medicina de Harvard, cualquier variable fisiológica, incluido el ritmo cardíaco, debería retornar a su estado estacionario "normal" después de haber sido perturbada. El principio de homeostasis sugiere que las variaciones del ritmo cardíaco son meras respuestas transitorias a las fluctuaciones del ambiente. Se podría postular razonablemente que, durante la enfermedad o el envejecimiento, el organismo es menos capaz de mantener un ritmo cardíaco constante en reposo, con lo que la magnitud de las variaciones del ritmo será mayor.

Pero cuando se procede a medir cuidadosamente las variaciones normales de ritmo cardíaco, latido a latido, y se las representa gráficamente a lo largo de un día, la imagen que pinta es muy otra. La representación gráfica de esta serie temporal aparece desgarrada, irregular y, a primera vista, completamente aleatoria. Pero cuando los datos sobre ritmo cardíaco se representan a diferentes escalas de tiempo, comienza a asomar una pauta. Si nos fijamos en unas horas de la serie temporal, encontramos fluctuaciones más rápidas cuyo alcance y secuencia tienen un aspecto semejante al de la gráfica original, correspondiente a un intervalo cronológico mayor. A escalas de tiempos más reducidas aún (minutos), encontramos variaciones más rápidas todavía, semejantes, lo mismo que antes, a la gráfica original. Las fluctuaciones latido a latido a diferentes escalas de tiempo dan la impresión de sibisemejanza característica de las ramas de un fractal geométrico. Este hallazgo induce a pensar que el mecanismo que controla el ritmo cardíaco podría ser intrínsecamente caótico. Dicho de otro modo, el ritmo cardíaco podría fluctuar considerablemente incluso en ausencia de estímulos externos fluctuantes, en lugar de relajarse a un estado homeostático, regular.

Para investigar si las variaciones, latido a latido, del ritmo cardíaco son efectivamente caóticas, o si son periódicas, podemos calcular el espectro de Fourier de la gráfica cronológica correspondiente al ritmo cardíaco. El espectro de Fourier de una forma de onda cualquiera (la de la gráfica cronológica) revela la pre-



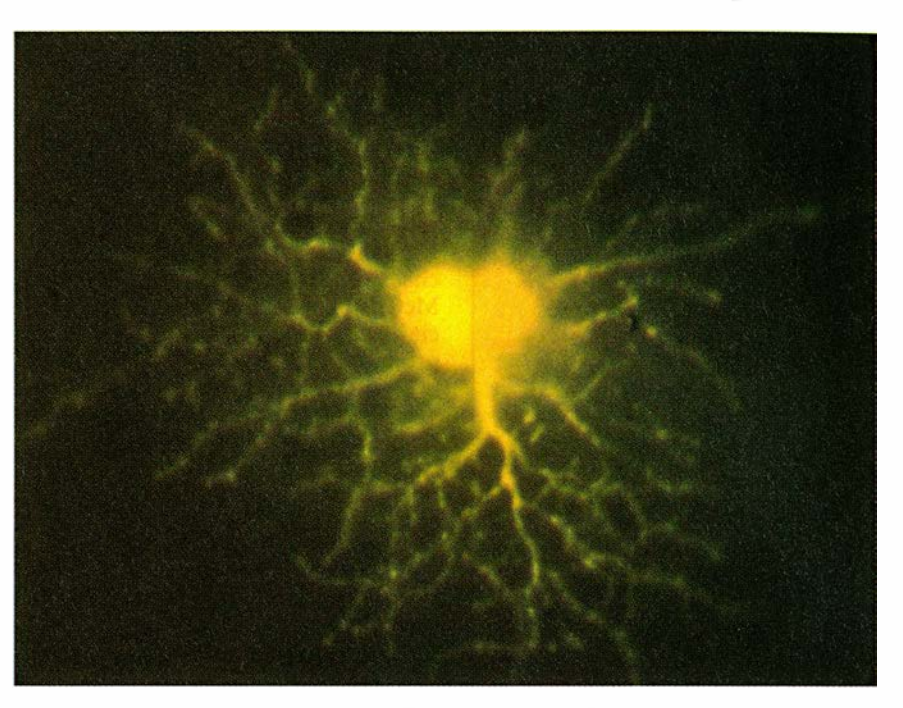

*4. RITMO CARDIACO, ofrecido en representaciones gráficas de series temporales* (izquierda), *espectros obtenidos por análisis armónico* (centro) *y trayectorias en el espacio de fases* (derecha). *Trece horas antes de un paro cardíaco, el ritmo era casi constante* (arriba), *como indican el espectro plano y la trayectoria en el espacio de fases, que sugiere un atractor puntual. Un ritmo cardíaco estudiado ocho días antes de una muerte cardíaca súbita* (centro) *muestra variación periódica; la trayectoria sugiere un ciclo límite ruidoso. Un ritmo cardíaco sano* (abajo) *ofrece un aspecto errático, posee un amplio espectro y una trayectoria que nos hace pensar en un atractor extraño.*

sencia de componentes periódicas. Si una serie cronológica mostrase un ritmo de exactamente un latido por segundo, el espectro mostraría una estrecha aguja a la frecuencia de un latido por segundo. Por otra parte, la gráfica de una serie cronológica correspondiente a pulsaciones caóticamente arrítmicas generaría un espectro que presentaría o bien amplias lomas o picos no bien definidos. Y la verdad es que el análisis espectral de la variabilidad del ritmo cardíaco normal muestra un amplio espectro que sugiere el caos.

Otro instrumento de análisis de la dinámica de un sistema complejo no lineal es la representación en el "espacio de fases". Esta técnica va siguiendo el rastro a los valores de las variables independientes que cambian en el transcurso del tiempo. El tipo y número de las variables independientes depende del sistema [véase "Caos", por James P. Crutchfield, J. Doyne Farmer, Norman H. Packard y Robert S. Shaw; INVESTIGACIÓN Y CIENCIA, febrero de 1987]. En muchos sistemas complejos resulta imposible identificar o medir fácilmente la totalidad de las variables independientes. En el caso de tales sistemas, las representaciones en el espacio de fases pueden efectuarse merced a la técnica de mapas de retardo. En el caso del mapa de retardo más sencillo, cada punto de la gráfica corresponde al valor de cierta variable en un instante dado, expresado en función del valor que toma esa misma variable tras un período de retardo fijo. Una serie de tales puntos, correspondientes a instantes sucesivos, esboza una curva o trayectoria que describe la evolución del sistema.

Para identificar el tipo de dinámica del sistema (caótica o periódica), se determinan las trayectorias correspondientes a una multitud de condiciones iniciales. Después, se busca un atractor, o sea, una región del espacio de fases que atraiga a las trayectorias. El atractor de tipo más sencillo es el punto fijo. Describe un sistema —como el de un péndulo amortiguado— que evoluciona siempre hacia un mismo y único estado. En la región del espacio de fases próximo a un atractor de tipo punto fijo, las trayectorias convergen hacia un punto.

El atractor de tipo inmediatamente más complicado es el ciclo límite. Corresponde a un sistema —como el de un péndulo ideal, sin rozamiento— que evoluciona siempre hacia un estado periódico. En el espacio de fases cercano al ciclo límite, las trayectorias siguen una senda regular, que es circular o elíptica.



*5. LAS NEURONAS constituyen ejemplos de estructura fractal. El cuerpo celular se ramifica en dendritas, las cuales se escinden, a su vez, en fibras más finas. Esta estructura puede tener relación con el caos en el sistema nervioso.*

Otros atractores son simplemente calificados de "extraños". Describen sistemas que no son ni estáticos ni periódicos. En el espacio de fases cercano a un atractor extraño, dos trayectorias que partieron de condiciones prácticamente idénticas divergerán a corto plazo y acabarán por ser muy diferentes a la larga. El sistema descrito por un atractor extraño es caótico.

Hemos analizado recientemente las representaciones de fase correspondientes al ritmo cardíaco normal. Lo que hallamos se asemejaba más a un atractor extraño que a un atractor periódico, característico de un proceso verdaderamente regular. Esta observación era otra indicación de que la dinámica del ritmo cardíaco normal podría ser caótica.

El mecanismo responsable del caos en la variabilidad de unos latidos a otros en el corazón sano reside, probablemente, en el sistema nervioso. El nodo senoideo (el marcapasos natural del corazón) recibe señales de la porción involuntaria (autónoma) del sistema nervioso. El sistema nervioso autónomo consta, a su vez, de dos grandes ramas: el sistema parasimpático y el sistema simpático. La estimulación parasimpática hace disminuir el ritmo de excitación de las células del nodo senoideo, mientras que la estimulación simpática produce el efecto contrario. La influencia de estas dos ramas tiene por efecto un incesante tira y afloja sobre el marcapasos. El resultado de esta disputa sin fin son las fluctuaciones en el ritmo cardíaco de los individuos sanos. Investigaciones recientes, entre las que figuran las debidas a Richard J. Cohen y sus colegas, en el Instituto de Tecnología de Massachusetts, han cuantificado la reducción de la variabilidad en los latidos que se produce tras un trasplante de corazón, proceso quirúrgico en el que las fibras nerviosas autónomas resultan seccionadas.

Pruebas recientes, obtenidas en diversos laboratorios, sugieren que el caos constituye una característica normal de otros componentes del sistema nervioso. Gottfried Mayer-Kress, del Laboratorio Nacional de Los Alamos, Paul E. Rapp, del Colegio Médico de Pennsylvania, y Agnès Babloyantz y Alain Destexhe, de la Universidad Libre de Bruselas, han analizado electroencefalogramas de individuos sanos y han encontrado pruebas de caos en el sistema nervioso. Otto E. Rössler y sus colegas, de la Universidad de Tubinga, han descubierto, asimismo, indicaciones de caos en componentes del sistema nervioso que son responsables de la secreción hormonal. Han analizado cambios temporales de los niveles hormonales en individuos sanos, hallando fluctuaciones aparentemente caóticas.

Otros investigadores han simulado hace poco las interacciones entre las células nerviosas, para ver cómo po-

dría surgir el caos. Walter J. Freeman, de la Universidad de California en Berkeley, ha puesto de manifiesto que es posible generar el caos en un modelo del sistema olfativo. El modelo se vale de un bucle de realimentación entre las "neuronas" y un retardo en los tiempos de respuesta. Con anterioridad, Leon Glass y Michael C. Mackey, de la Universidad McGill, han reconocido la importancia de los retardos temporales en la producción de caos.

¿Por qué habrían de exhibir una dinámica caótica el ritmo cardíaco y otros sistemas controlados por el sistema nervioso? Tales dinámicas pueden ofrecer numerosas ventajas funcionales. Los sistemas caóticos operan bajo una amplia gama de condiciones y son, por consiguiente, adaptables y flexibles. Esta plasticidad permite a los sistemas habérselas con las exigencias de un ambiente cambiante e impredictible.

Muchas patologías exhiben un comportamiento cada vez más periódico y una pérdida de variabilidad. Las primeras indicaciones de que incluso el corazón agonizante puede comportarse periódicamente procedieron del análisis armónico de formas de onda electrocardiográficas de Fourier durante la taquicardia ventricular o la fibrilación ventricular, que son los ritmos cardíacos sumamente rápidos que con gran frecuencia provocan el paro cardíaco. A mediados del pasado decenio, Raymond E. Ideker y sus colegas, de la facultad de medicina de la Universidad de Duke, registraron las formas de onda asociadas con la fibrilación ventricular desde los estratos más profundos del corazón de perro. Hallaron que la actividad fibriladora en el seno del corazón era un proceso de carácter mucho más periódico de lo que se había pensado hasta entonces.

En 1988, dos de los autores (Goldberger y Rigney) efectuaron un estudio retrospectivo de los electrocardiogramas de ambulatorio de personas que sufrían enfermedades cardíacas graves. Descubrimos que la pauta de latidos de esos pacientes solía hacerse menos variable de lo normal en un período que iba desde minutos hasta meses antes de la muerte cardíaca súbita. En algunos casos, se reducía la variabilidad global de latido a latido; en otros, aparecían y desaparecían abruptamente oscilaciones periódicas en el ritmo cardíaco.

De forma un tanto similar, el sistema nervioso puede mostrar la pérdida de variabilidad y la aparición de periodicidades patológicas en la epilepsia, la enfermedad de Parkinson y la depresión maníaca. Y mientras que se ha informado de que, en condiciones normales, los recuentos de glóbulos blancos dan cifras caóticamente variables día a día en individuos sanos, en ciertos casos de leucemia, dicha cifra oscila periódicamente.

Las regularidades periódicas en la enfermedad y el comportamiento aparentemente caótico que se observa en la salud no implican que todas las patologías se encuentren asociadas a una creciente regularidad. En algunas arritmias cardíacas, el pulso varía de forma tan errática que el individuo puede quejarse de "palpitaciones". Algunos de estos acontecimientos representan oscilaciones que parecen irregulares, aunque revelan su carácter periódico real en cuanto se las analiza con detenimiento. En otras arritmias, los latidos son realmente impredeciblemente erráticos. Sin embargo, no se ha demostrado que ninguna de estas patologías irregulares sea representación de caos no lineal, por más que el pulso puede dar una impresión enteramente "caótica" en el sentido coloquial.

La fisiología puede demostrar ser uno de los más ricos laboratorios para el estudio de los fractales y el caos, amén de otros tipos de dinámicas no lineales. Los fisiólogos tienen necesidad de lograr una comprensión más completa de la forma en que los procesos de desarrollo conducen a la construcción de arquitecturas fractales y de por qué los procesos dinámicos del organismo generan un aparente caos. En años no lejanos, los estudios de los fractales y el caos en fisiología pueden proporcionar medios más sensibles de caracterizar la disfunción producida por el envejecimiento, la enfermedad y la toxicidad de las drogas.

BIBLIOGRAFIA COMPLEMENTARIA

AN ESSAY ON THE IMPORTANCE OF BEING NONLINEAR. B. J. West en *Lecture Notes in Biomathematics 62*. Dirigido por S. Levine. Springer-Verlag, 1985.

FRACTALS IN PHYSIOLOGY AND MEDICINE. Ary L. Goldberger y Bruce J. West en *Yale Journal of Biology and Medicine*, vol. 60, págs. 421-435; 1987.

PHYSIOLOGY IN FRACTAL DIMENSIONS. Bruce J. West y Ary L. Goldberger en *American Scientist*, vol. 75, n.º 4, págs. 354-365; julio-agosto de 1987.

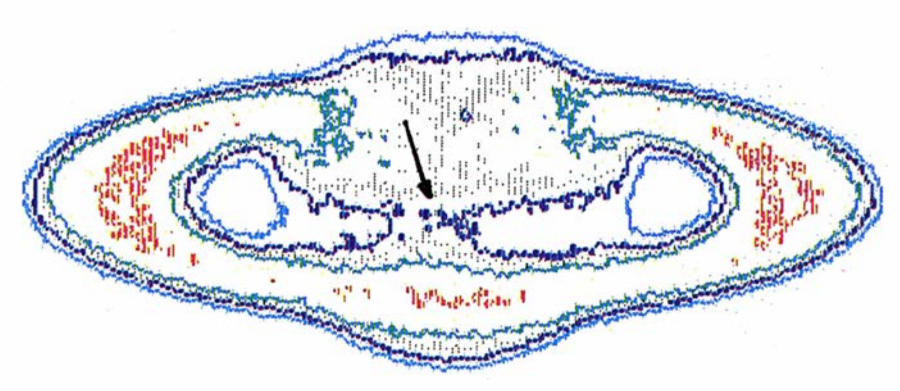
NONLINEAR DYNAMICS IN SUDDEN CARDIAC DEATH SYNDROME: HEARTRATE OSCILLATIONS AND BIFURCATIONS. A. L. Goldberger, D. R. Rigney, J. Mietus, E. M. Antman y S. Greenwald en *Experientia*, vol. 44, págs. 983-987; 1988.

# Ciencia y sociedad

## *Grandes manchas blancas de Saturno*

Saturno es un planeta deslumbrante cuando se visualiza a través del telescopio. El observador se ve inmediatamente atraído por el impresionante sistema de anillos que rodea su plano ecuatorial. Sólo en raras ocasiones, aproximadamente una vez cada quince años, éstos se muestran de canto, desapareciendo de la visión telescópica debido a su fina estrechez.

Por contra, si dirigimos la mirada hacia el disco planetario, apenas si se perciben tenues rasgos de las estructuras nubosas de la atmósfera: algunas bandas (franjas oscuras) y zonas (franjas más claras) discurren paralelas al ecuador. Comparado con Júpiter, en donde la variedad de detalles cambiantes abunda (vórtices ovales, entre ellos la Mancha Roja, festones ecuatoriales, etcétera), la atmósfera de Saturno aparece decepcionante. La ausencia total de signos de actividad meteorológica es la nota. Los registros históricos apenas muestran la presencia de un par de docenas de detalles que esporádicamente fueron detectados en diferentes latitudes. Su seguimiento permitió medir la rotación diferencial de la atmósfera, es decir, la dependencia de la intensidad de los vientos dominantes que circulan paralelos al ecuador con la latitud. Sin embargo, de entre todos los detalles destacan por su excepcionalidad, virulencia, tamaño y grado de actividad, unas estructuras que hemos bautizado con el nombre de "Grandes Manchas Blancas" (GWS de sus iniciales en inglés), y que llegan incluso a superar a las estructuras más activas observadas en la atmósfera de Júpiter.



*1. Gran Mancha Blanca de 1933, fotografiada durante el mes de agosto desde el Observatorio Lowell. Debajo, contorno de isófotas de la imagen anterior efectuado por el autor con el microfotómetro* PDS *del Instituto de Optica de París.*

Tenemos constancia de que desde el último tercio del siglo pasado se han producido al menos cuatro de tales fenómenos. La primera gran mancha fue detectada por el astrónomo norteamericano Asaph Hall en 1876 cerca del ecuador del planeta. Corresponde al también norteamericano E. E. Barnard el descubrimiento de la segunda, que apareció en las latitudes medias del hemisferio Norte, y que entre otros también fuera estudiada por el astrónomo catalán José Comas Solá empleando su telescopio de 156 mm de abertura. El tercer evento fue observado por el astrónomo aficionado inglés W. T. Hay en 1933, evolucionando sobre la región ecuatorial del planeta. Esta mancha fue una de las mejor analizadas ya que fue seguida por un gran número de astrónomos que emplearon diversas técnicas. El cuarto y último episodio conocido fue registrado por el aficionado sudafricano J. H. Botham en 1960. Se trató de la mancha más boreal, pues emergió en la latitud 58° del hemisferio Norte.

Cuando una de tales manchas aparece, destaca por su brillo, resaltando por contraste con respecto al fondo amarillento del disco; sobresale también por su tamaño, de unos 20.000 kilómetros en promedio (aproximadamente un tercio del radio del planeta). Su forma, en un comienzo redondeada y compacta (bordes bien definidos), comienza a los pocos días a expandirse zonalmente (es decir, en dirección de los paralelos) y en el sentido contrario al del avance de la mancha, hasta alcanzar en una semana los 100.000 km de extensión. Esta expansión va acompañada de una pérdida paulatina de brillo.

Para las manchas de 1933 y 1960 hemos medido velocidades del movimiento expansivo de 30 a 40 metros por segundo. En el caso de la mancha de 1960 se produjo simultáneamente una expansión en dirección Norte, con velocidad de 3 metros por segundo, llegando a alcanzar la latitud 80°. Al cabo de un mes de comenzar el evento, la actividad se ha extendido a todas las longitudes del planeta –lo rodea por completo–, ocupando principalmente toda la zona en donde emergió, de unos 10° de amplitud en latitud. Esta franja presenta un aspecto parcheado con regiones claras alternando con otras oscuras.

Las GWS no son, en sí mismas, de-

talles nubosos estáticos con respecto al interior de Saturno, sino que se desplazan de acuerdo con los vientos zonales dominantes. Así, las manchas ecuatoriales de 1876 y 1933 se movían en promedio hacia el Este con velocidades de unos 400 metros por segundo. La de 1903, a 36° N, lo hizo a unos 20 metros por segundo; la boreal de 1960, a 58° N, procedió a 4 metros por segundo, si bien hacia el Oeste. Estos valores están en buen acuerdo con los medidos en esas latitudes por los vehículos espaciales *Voyager 1* y *2* durante su encuentro con el planeta en 1980 y 1981. Esto parece indicar la persistencia en el tiempo de las corrientes de vientos en su magnitud y posición en el planeta.

Nuestro análisis de las fotografías tomadas bajo diversos filtros, desde el ultravioleta al rojo, de las manchas de 1933 (fueron varios los focos de actividad observados a lo largo de un par de meses), así como las medidas de la proporción de luz reflejada polarizada en la del año 1960, sugieren que las GWS son detalles altos de la atmósfera; es decir, la cima nubosa de las manchas se sitúa por encima del nivel medio de la capa superior de nubes de Saturno.

El hecho de que sean detalles elevados, unido a la naturaleza eruptiva observada durante su aparición y al gran dinamismo mostrado en su evolución posterior, nos ha llevado a indagar en la posibilidad de que las manchas sean la consecuencia de un proceso de inestabilidad convectiva. Formulamos para ello un modelo según el cual las GWS se originan por convección húmeda en las regiones más calientes e inferiores de la atmósfera (presiones del orden de unos 10 bar), en donde, debido a la liberación del calor latente producido principalmente por la condensación del vapor de agua, tiene lugar el ascenso y propulsión de enormes masas de gas. Cuando éstas alcanzan alturas de unos 150 km (presiones del orden de 1 bar) por encima de aquel nivel, el gas amoníaco presente en la atmósfera se condensa, debido a las bajas temperaturas allí reinantes, y forma pequeños cristalitos que en su conjunto constituyen el techo de las manchas. El mecanismo semeja así lo que acontece en la formación de los cumulonimbos de la atmósfera terrestre que dan origen a violentas tormentas, aunque en una escala unas cien veces mayor. Con una excepción: en la Tierra, la densidad del aire saturado con vapor de agua es menor que la del aire seco, lo que favorece el ascenso. Lo contrario es válido en los planetas gigantes, en donde la atmósfera seca contiene gases muy ligeros: hidrógeno (en una proporción del 90 por ciento) y helio (casi todo el resto). Por ello es de suponer la implicación de mecanismos supletorios que ayuden, al menos inicialmente, la propulsión vertical de las masas de gas.

Una posibilidad que hemos apuntado es la siguiente: el mecanismo actuante es de origen dinámico, de naturaleza ondulatoria, que corresponda a la propagación de "ondas de gravedad atmosférica" que favorezcan la convergencia del aire en niveles inferiores. Se sabe que un proceso de este tipo opera eficazmente en la llamada zona de convergencia intertropical en las latitudes ecuatoriales de nuestra propia atmósfera.

Los cuatro eventos registrados hasta la fecha acontecieron, como ya dijimos, en latitudes dispares, desde el ecuador hasta las regiones subpolares, y todas ellas en el hemisferio Norte. Debido principalmente a la inclinación de unos 27° del eje de rotación del planeta respecto del plano orbital, la insolación varía sobre Saturno de forma estacional. Pues bien, curiosamente las manchas aparecieron en aquellas latitudes en donde la insolación era máxima entonces; todas ellas, cuando el planeta se encontraba cerca de una misma posición orbital. Esto implica una recurrencia de una por cada año de Saturno, cuya duración es de unos 29,5 años terrestres. Parece, por tanto, que el caldeo solar puede influir en la aparición de las manchas.

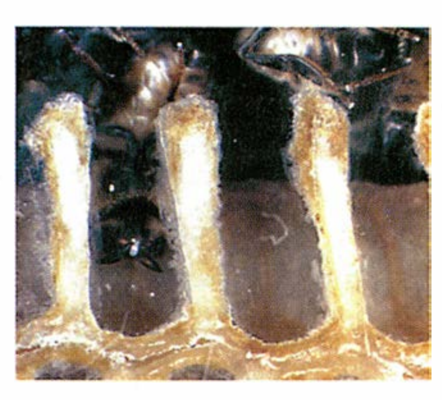
La periodicidad apuntada debe ser tomada, sin embargo, con precaución, ya que se produce un sesgo en la observación de uno de los hemisferios del planeta durante quince años, a causa de la inclinación del eje de rotación y el consiguiente efecto de bloqueo en la visibilidad que originan los anillos. En cualquier caso, Saturno se encontrará en las proximidades del punto orbital de interés durante este y el próximo año. Sería conveniente vigilar el planeta para ver si se produce una de estas raras erupciones de nubes blancas. El tiempo lo dirá. (A. Sánchez Lavega.)

## *Abejas policía*

Con frecuencia se ha admirado a las obreras de las abejas melíferas por su civismo desinteresado. Mientras que la abeja reina y los zánganos sólo viven para reproducirse, las obreras realizan las tareas domésticas necesarias para mantener la colmena zumbando: la recolección del néctar, la construcción de los panales y la alimentación de la progenie de la coronada. Pero experimentos recientes revelan que las obreras no son tan altruistas como parecen: producen la ilusión de lealtad voluntaria a su reina mediante la supresión brutal de los intentos de las demás obreras de competir con la reina en la producción de descendientes.







*2. Abeja obrera que encuentra y luego se come un huevo puesto por otra obrera, en estas fotografías realizadas por Francis L. W. Ratnieks.*

Un repaso a la genealogía apiar ayudará a iluminar este oscuro secreto de familia. Típicamente, la reina es la madre de todas las abejas de una colmena; después de aparearse con hasta 20 zánganos de otras colonias, engendra hijas que se convierten en obreras. Las obreras de un mismo pa-

dre son hermanas completas, pero la mayoría son sólo medias hermanas las unas de las otras. La reina también pone huevos no fertilizados que se convierten en zánganos, que son hermanos de todas las obreras de la colonia.

Que las obreras no puedan aparearse, no significa que sean del todo estériles. Cuando una reina muere, a menudo las obreras ponen huevos no fertilizados que, como los de la reina, dan origen a zánganos masculinos. Pero las obreras raramente se reproducen mientras reina su soberana. El pasado otoño, P. Kirk Visscher, de la Universidad de California en Riverside, informó que las obreras producen aproximadamente uno de cada 1000 de los zánganos de la colmena.

¿Por qué razón las obreras no producen más descendientes masculinos cuando vive la reina? Después de todo, el hijo de una obrera porta más de sus genes que sus hermanos, los zánganos producidos por la reina. Según la moderna teoría evolutiva, una obrera aumentaría su eficacia biológica (o capacidad de propagar sus genes) produciendo sus propios huevos en adición a los de la reina, o en lugar de ellos.

Por otra parte, la eficacia de una obrera típica disminuiría si los hijos de la reina fueran suplantados por los de otras obreras, la mayoría de los cuales sólo serían medios sobrinos de la obrera. A partir de estas premisas, Francis L. W. Ratnieks, de la Universidad de California en Berkeley, propuso hace dos años que las obreras bloquean los intentos de reproducirse de las demás obreras, mediante una "vigilancia policial" mutua.

Ratnieks y Visscher se asociaron para someter la hipótesis a prueba. Introdujeron en una colmena huevos no fertilizados puestos por obreras y por la reina. Como sospechaban, las obreras de la colmena olieron con rapidez y devoraron casi todos los huevos de las obreras, al tiempo que dejaban incólumes la mayoría de los huevos de la reina. ¿Ocurre con frecuencia esta vigilancia policial, o bien a la mayoría de las obreras se les impide que pongan? "Esto es precisamente lo que estamos estudiando ahora", dice Visscher.

En un comentario que acompaña el artículo de Ratnieks y Visscher en *Nature*, Jon Seger, de la Universidad de Utah, compara la sociedad de abejas melíferas con el estado totalitario que George Orwell describió en *Mil novecientos ochenta y cuatro*. Ratnieks sospecha que el fundamento de la cooperación entre otros insectos muy organizados, como las hormigas, puede ser la coacción. "Quizá los insectos se parezcan más a nosotros de lo que creíamos", señala.

## *¿Quien tiene la banca?*

De la misma manera que el banquero de una partida de póquer puede disponer los 52 naipes en casi 2,6 millones de manos distintas, enzimas sin identificar del sistema inmune combinan de formas diversas unos cuantos fragmentos de ADN para crear los genes de miles de millones de anticuerpos diferentes. Este truco de barajar los genes es decisivo para las defensas naturales del cuerpo; si los linfocitos, células del sistema inmune, tuvieran que transportar la dotación completa de genes para todos los anticuerpos que se precisan en la protección del organismo contra los invasores, no quedaría ADN para codificar nada más. Dos grupos en los Estados Unidos y el Japón acaban de identificar los genes que determinan dos proteínas distintas, una de las cuales puede ser la recombinasa, el "banquero" enzimático de esta partida inmunológica que hace mucho tiempo que se buscaba.

La búsqueda de esta escurridiza entidad empezó en 1976, cuando Susumu Tonegawa y Nobumichi Hozumi, a la sazón en el Instituto de Inmunología de Basilea, demostraron que los genes que codificaban las proteínas de los anticuerpos se ensamblaban al azar a partir de segmentos de genes distintos: variables (*V*), de unión (*J*) y, a veces, de diversidad (*D*) [véase "Moléculas del sistema inmunitario", de Susumu Tonegawa; INVESTIGACIÓN Y CIENCIA, diciembre de 1985]. La enzima hipotética que podía reconocer y empalmar tales segmentos se denominó recombinasa *V*(*D*)*J*.

Durante muchos años, el mejor candidato para la recombinasa *V*(*D*)*J* pareció ser el producto de un gen denominado *scid*, siglas de la traducción inglesa de "inmunodeficiencia grave combinada". Los ratones mutantes que manifiestan *scid* son incapaces de ensamblar adecuadamente los genes de la inmunoglobulina. Aunque no se ha logrado identificar el gen *scid*, se sabe que se encuentra en el cromosoma 16.

En la primavera de 1986, David G. Schatz y David Baltimore, del Instituto Whitehead de Investigación Biomédica, empezaron a acotar un gen que se hallaba implicado de manera principal en la recombinación *V*(*D*)*J*. Demostraron que ADN procedente de linfocitos podía disparar la recombinación *V*(*D*)*J* en fibroblastos, una clase de células del tejido conjuntivo que no acostumbra recombinar los segmentos. Más recientemente, en *Cell*, Schatz, Baltimore y su colega Marjorie A. Oettinger anunciaron que habían aislado el gen único responsable de esta actividad; lo bautizaron "gen activador de la recombinación", RAG-1. Pruebas efectuadas por Schatz, Oettinger y Baltimore demostraron que RAG-1 no era *scid*: los dos genes se alojaban en cromosomas distintos.

¿Es RAG-1 el gen para la recombinasa, o bien capitanea otras enzimas subordinadas de la recombinasa? Schatz admite que las pruebas de que se dispone son ambiguas y que "ésta es la cuestión crítica que guiará nuestra investigación futura". Sin embargo, señala que ciertas pruebas circunstanciales, recabadas en distintos frentes, sugieren que la proteína RAG-1 es la recombinasa. La evolución ha conservado RAG-1 en animales tan lejanamente emparentados como ranas, caballos, gallinas y seres humanos, según cabía esperar de enzima tan importante. Se expresa naturalmente sólo en los linfocitos en desarrollo y únicamente durante los estadios en los que tiene lugar recombinación. Lo que es quizá más importante, el RAG-1 no parece conferir a los fibroblastos ninguna de las propiedades de los linfocitos, excepto la recombinación *V*(*D*)*J*; cabe esperar que una proteína de desarrollo o reguladora desencadene, asimismo, otros cambios.

Otro candidato posible para la recombinasa es el producto del gen RBP-2, lo que recientemente anunció en *Nature* un grupo dirigido por Tasuku Honjo, de la Universidad de Kyoto. La proteína de RBP-2 se une de forma específica a la "secuencia de reconocimiento" que flanquea los segmentos del gen *J* antes de la recombinación; la proteína posee asimismo grandes semajanzas con la clase de proteínas denominadas integrasas, que empalman nuevo ADN en los cromosomas bacterianos. Todavía no se ha demostrado, sin embargo, que el RBP-2 desempeñe un papel en la recombinación *V*(*D*)*J*, lo que hace claramente el RAG-1.

Según Frederick W. Alt, inmunólogo de la Universidad de Columbia, todavía no está claro de qué modo encajan RAG-1, RBP-2 y *scid* en un cuadro coherente de actividad recombinasa. La proteína RAG-1 podría ser la recombinasa principal, y las proteínas RBP-2 y *scid* actuar como

factores accesorios necesarios. Alternativamente, puede que RAG-1 sea sólo un conmutador maestro que pone en funcionamiento las proteínas recombinasas de RBP-2 y de otros genes.

## *Potpourri astronómico*

El pasado mes de enero se reunieron en Washington D.C. mil setecientos astrónomos con ocasión del Congreso anual de la Asociación Astronómica Americana, para presentar e intercambiar puntos de vista sobre los últimos descubrimientos de la investigación cosmológica. Con toda certeza, la mayor sorpresa del encuentro fue la noticia de que el *Explorador de la radiación cósmica de fondo* (COBE), ingenio lanzado el 18 de noviembre de 1989, ha obtenido ya mediciones precisas de las penetrantes emisiones infrarrojas y de microondas que constituyen, así se cree, los restos de la gran explosión (el "big bang") que significó el comienzo de nuestro universo.

Pero lo interesante del caso es, justamente, lo que no se detecta, en franco contraste con las observaciones realizadas en 1987 que mostraban complicadas irregularidades. El espectro de la radiación de fondo no muestra ninguna desviación del espectro que produce un objeto radiante ideal. El COBE no ha encontrado tampoco signos de trozos o explosiones del universo primitivo que expliquen las agrupaciones de estructuras que los astrónomos ven hoy. Conforme el COBE vaya suministrando nuevos datos, éstos mejorarán progresivamente la calidad de la información. Si no se observan cúmulos en ese nivel, los cosmólogos tendrán que desechar muchas teorías actuales sobre la formación de galaxias.

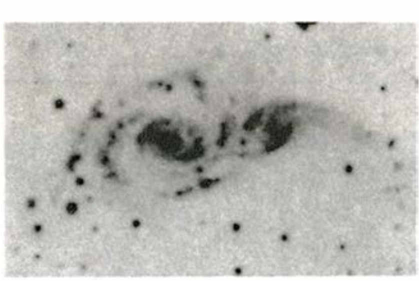
Por otra parte, continúan acumulándose las pruebas que perfilan un universo preñado de enormes y complejas estructuras. Alan M. Dressler, de la Carnegie Institution de Washington, y Sandra M. Faber, de la Universidad de California en Santa Cruz, han descrito el movimiento de las galaxias alrededor del "Gran Atractor" —una extraordinaria acumulación de galaxias, que cubre un tercio del firmamento del hemisferio Sur; atrae, hacia sí, la Vía Láctea y todas las galaxias cercanas— y han demostrado que se trata de una estructura real con un extremo observable.

Ed Shaya, de la Universidad de Columbia, y R. Brent Tully, de la Universidad de Hawai, quienes han trabajado en el cálculo de la masa del Gran Atractor para determinar si podría explicar, por sí mismo, el movimiento de la Vía Láctea en aquella dirección, han llegado a la conclusión de que las galaxias observadas dan cuenta sólo de dos tercios de su movimiento. Proponen, en consecuencia, la existencia de otro centro mayor que el anterior, incluso más distante: un "Ingente Atractor".

T.J. Broadhurst y sus colegas, de la Universidad de Durham, presentaron un descubrimiento muy sugestivo, aunque teórico: la densidad de galaxias varía de manera regular por casi todo el universo visible, con máximos que se repiten cada 400 millones de años-luz. El mecanismo que puede generar estas enormes estructuras se desconoce.

Otro sorprendente hallazgo se refiere al descubrimiento de un cuásar doble. La mayoría de los cuásares dobles se cree que son ilusiones ópticas producidas por la acción de galaxias





*3. Galaxias en interacción, capaces de producir galaxias "oculares"* (arriba). *Estas podrían acabar en una estructura segmentada, como la observada hace poco en la Vía Láctea. En choques más violentos podría enviarse materia hacia agujeros negros del centro de las galaxias y producir cuásares dobles* (simulación inferior).

remotas y dotadas de gran masa sobre la luz emitida por estos objetos cósmicos. Un equipo dirigido por Georges Meylan, del Instituto Científico del Telescopio Espacial, asegura haber descubierto el primer caso conocido de dos cuásares dobles reales. Los dos objetos están, tal parece, a la misma distancia de la Tierra —unos doce mil millones de años luz—, lo cual implica que uno dista de otro cien mil años luz, menos que el diámetro de la Vía Láctea.

Tamaña proximidad abona la idea de una interacción gravitatoria mutua. Las teorías actuales suponen que la cercanía entre cierto tipo de galaxias podría incoar la formación de cuásares incipientes en los núcleos de las mismas. Meylan especula sobre la posibilidad de que estos dos cuásares estén experimentando un proceso similar y que nosotros podamos ser testigos del nacimiento de un apareamiento de cuásares.

A otro nivel, a la vieja imagen de la Vía Láctea que nos la representa con un núcleo esférico de estrellas del que salen unos brazos, asimismo estelares, le han dado la vuelta Leo Blitz, de la Universidad de Maryland en College Park, y David N. Spergel, de la Universidad de Princeton; en su opinión, la región interna posee forma de pelota de rugby, con lo que la Vía Láctea no sería una galaxia espiral, sino una configuración blanda de "espiral segmentada" cuyos brazos saldrían de una barra estelar, y no de una esfera.

¿Cómo justificar una estructura de este estilo? Debra Meloy Elmegreen, del Vassar College, que ha estudiado una clase muy peculiar de galaxias interactivas a las que ha calificado de "oculares" por su semejanza con el globo ocular, afirma que estas galaxias son de evolución rápida, en términos cósmicos (quizás unos doscientos millones de años), y que acaban convirtiéndose en galaxias espirales segmentadas. La Vía Láctea tiene dos galaxias próximas, las Nubes de Magallanes, con las que podría haber interactuado en el pasado, aunque, según Elmegreen, sus simulaciones de estas interacciones en un ordenador no han podido crear una estructura de pelota de rugby como la observada.

## *¿Cómo construir un cable superconductor?*

Los cables flexibles formados con superconductores calientes podrían permitir aplicaciones tan extraordinarias como los motores superconductores. Sin embargo es muy difícil conseguirlos. Investigadores del Instituto de Tecnología de Georgia en Atlanta han desarrollado una técnica que posibilitará cubrir fibras con materiales superconductores y formar cables con estas fibras.

W. Jack Lackey y sus colegas han modificado una técnica de cubrimiento común: la deposición química de vapor (DQV). En esta técnica convencional, se calienta una substancia hasta vaporizarla; el gas fluye entonces hacia otra cámara donde se condensa sobre un sustrato.

Dado que los reactivos para materiales superconductores formados por itrio, bario y cobre tienen presiones de vapor pequeñas, las correspondientes tasas de deposición resultan ser de 1 micrometro por hora. Para los superconductores, la técnica DQV resulta harto complicada, ya que cada material debe vaporizarse por separado y, en cambio, han de introducirse en la cámara de condensación todos juntos. Esto requiere el control simultáneo de la presión, de la temperatura y de la velocidad de flujo de los tres gases.

Cansados de trabajar con la técnica DQV, a Lackey se le ocurrió la idea de emplear los elementos triturados. Moliendo los tres constituyentes muy finamente y mezclándolos en proporciones regulares con argón, creó un sistema eficaz de transporte de los materiales hacia la cámara del sustrato. Al calentarse este recinto, se vaporiza el polvillo y se deposita en el sustrato.

Los investigadores afirman que han depositado capas de materiales superconductores de gran calidad con tasas de 200 micrometros por hora. Aunque hasta ahora sólo han trabajado con sustratos planos, Lackey ha diseñado un sistema DQV dotado de poleas que le permitirá cubrir fibras a medida que se van transportando a través del dispositivo.

Afirma también que en las primeras muestras obtenidas han circulado intensidades de 40.000 ampère por centímetro cuadrado, suficientes para algunas aplicaciones. Los investigadores esperan producir cables largos que transporten corrientes muy altas.

# Ciencia y empresa

## *Empresas a la greña*

El año pasado, la empresa E. I. DuPont de Nemours pinchó en hueso. Había empezado a trabajar en la técnica de reacción en cadena de polimerasas (RCP), método revolucionario para multiplicar segmentos de ADN que había sido desarrollado antes por los investigadores de Cetus. DuPont trató, en vano, de obtener de Cetus la licencia para la técnica de RCP. Pero no se arredró: demandó a Cetus, afirmando que las patentes RCP de Cetus no eran válidas.

La pelea DuPont-Cetus fue uno más de los miles de pleitos sobre propiedad intelectual incoados en los últimos doce meses. La mayoría de los abogados mercantiles y expertos en patentes creen que éstas han subido su valor de cinco años a hoy. El mundo empresarial comienza a jugar fuerte ante los tribunales. Algunas compañías esperan cobrar regalías a los competidores; otras confían lisa y llanamente en eliminarlos del mercado. Esas tendencias se dejan sentir, sobre todo, en las empresas jóvenes, que están poniendo a prueba, por primera vez, los límites de la protección de la propiedad.

El interés por los derechos de propiedad intelectual se avivó hacia fines del decenio de 1970, cuando los EE.UU. comenzaron a preocuparse por su déficit comercial. Ante la competencia interior y exterior, las empresas hicieron un inventario detallado de sus bienes, patentes incluidas.

El resultado de los litigios, no obstante, era impredecible. Cada tribunal había de juzgar patentes lo mismo que demandas, y hubo una ventisca de decisiones contradictorias. El albur de la patente en disputa dependía del juez que a una empresa le tocara en suerte.

Bajo la administración Reagan, sin embargo, se empezó a considerar las patentes una herramienta "pro-competitiva". Por esas mismas fechas, el Congreso inició una labor legisladora sobre patentes cuya normativa no había cambiado desde los años cincuenta. Se han aprobado ya más de una docena de leyes. La más importante de ellas, según parecer unánime de los abogados, ha creado un tribunal especial para ese campo mercantil.

El tribunal comenzó rápidamente a desarrollar un conjunto más coherente de precedentes. Puso también de manifiesto el valor de las patentes con una decisión de 1985 que dirimía la batalla entre Polaroid y Eastman Kodak. Confirmó las patentes de Polaroid sobre cámaras para fotos al instante, y dejó a Kodak fuera de ese negocio.

El Congreso también reforzó la protección de patentes contra infractores extranjeros con la Ley de Comercio de 1988. Durante mucho tiempo, había estado prohibido que las empresas extranjeras vendieran sin licencia productos patentados en los EE.UU. Pero esta protección no cubría procesos patentados, a menos que los propietarios norteamericanos de la patente pudieran demostrar el perjuicio económico sufrido. La ley de comercio levantó esta restricción, proporcionando así la misma protección a las patentes de productos que a las de procesos.

Aun cuando los derechos de patentes se han reforzado, las industrias de punta, como las de soporte lógico informático y las de biotecnología, han empezado a medir sus fuerzas. Hasta ahora, los derechos de autor y las patentes no han perdido vigor. Así, no hace mucho algunos tribunales empezaron a fallar que la patente protegía también la "cadencia" o "morfología" de los programas informáticos, que suelen ir envueltos en elementos gráficos. Esas primeras decisiones desencadenaron un alud de pleitos, entre ellos dos interpuestos por Apple Computer contra Microsoft y Hewlett-Packard.

Según la leyenda industrial, el fundador de Apple, Steven P. Jobs, vio los primeros iconos de ordenador cuando visitaba el centro de investigación de Xerox. El ulterior uso por parte de Apple de los iconos en su soporte lógico para ordenadores personales le hizo apuntarse un enorme éxito; mientras tanto, Xerox se debatía torpemente en vanos intentos por entrar en el negocio.

Ahora que los tribunales están protegiendo los derechos de autor, Xerox tiene la esperanza de recibir un pago atrasado por su trabajo pionero. A principios de diciembre, Xerox pidió a Apple que le abonara los derechos de licencia. Apple se negó y Xerox interpuso un recurso para que los derechos de autor de Apple se borraran del registro federal. Además, Xerox pide más de 150 millones de dólares por daños y perjuicios.

La ola de pleitos ha recalado en la playa de la biotecnología. Se divisan nuevos mercados. ¿Cuánto valdría un específico recombinante para el SIDA? Algunas empresas han apostado su futuro a uno o dos productos, y muestran un obvio interés máximo por convertirse en los únicos proveedores.

Cualquier sospecha se lleva al juez, al tiempo que se demanda del Congreso una legislación más dura sobre la propiedad intelectual. Tenemos un caso paradigmático en la prolongada disputa entre Amgen y Genetics Institute (GI) sobre la eritropoyetina (EPO), proteína que regula la tasa de formación de los glóbulos rojos. Amgen va por delante en la investigación científica, pero un fallo reciente que reconocía ciertos derechos sobre EPO a ambas compañías ha dejado a Amgen en desventaja. El caso se ha complicado por la unión de GI con la firma japonesa Chugai Pharmaceutical. A raíz de ello, Amgen ha elevado su caso al Congreso, arguyendo que las leyes de patentes existentes están sesgadas en contra de las propias empresas norteamericanas.

Con apoyo por parte de otras empresas de biotecnología, en particular Genentech, se espera que la Cámara tome en consideración al menos dos proyectos de ley de propiedad intelectual. Uno tratará de facilitar la obtención de patentes de proceso para trabajos basados en técnicas recombinantes; el otro ampliará la protección a los productos biológicos intermedios usados en tecnologías recombinantes.

Mientras tanto, la industria biotecnológica aguarda el desenlace, sobre RCP, entre DuPont y Cetus. Es una lucha por una parte del mercado. A principios de 1989, Cetus concedió una licencia exclusiva a Hoffmann-La Roche para usar RCP con fines de diagnosis. Roche desea convertirse en uno de los primeros en el negocio de equipos de diagnóstico, y Cetus no quiere perder su lucro en el futuro del mismo.

DuPont, por su parte, no se aviene a quedar fuera de un negocio que al-

gunos predicen valdrá hasta 1000 millones de dólares dentro de un decenio. Cuando DuPont desenterró dos trabajos publicados hace más de 10 años que, según afirma, "anticipaban claramente" la técnica de reacción en cadena de polimerarsas denunció a Cetus por patentes inválidas.

DuPont y Cetus están de acuerdo en un punto: RCP vale las costas del pleito. El litigio puede fácilmente costar a cada compañía dos millones de dólares por año y puede durar más de dos años y medio. (Y ello sin incluir el gasto de las investigaciones; DuPont, por ejemplo, proyecta invertir más de 15 millones de dólares en la técnica RCP durante los próximos tres años.)

## *No sólo una cara bonita*

En un cuarto oscuro, con acordes de Wagner como fondo, Michael Barnsley está preparando una presentación. La pantalla de un ordenador Sun queda oscura durante unos 60 segundos; de repente, aparece la imagen en color de la cara de una mujer llenando la pantalla.

Esa imagen, dice Barnsley, se construyó a partir de sólo 5800 bytes de memoria del ordenador, un espacio asombrosamente pequeño para almacenar tal imagen. ¿Cómo? Barnsley dice que ha descubierto una técnica para comprimir automáticamente los datos en proporción de 500 a uno o más.

En una edad de copiosa información, hallar espacio en la memoria del ordenador para almacenar todos esos datos puede resultar difícil. La versión digital sin comprimir de la imagen de la mujer ocuparía unos tres millones de bytes de memoria; las técnicas de compresión convencional la reducen a unos cientos de miles de bytes. Añadir color e imágenes de vídeo en los ordenadores exige, sin embargo, una compresión todavía mayor.

Entre los que trabajan duro por hallar nuevas vías de compresión se encuentran Barnsley y Alan D. Sloan, matemáticos procedentes del Instituto de Tecnología de Georgia.

Hace varios años, Barnsley desarrolló algoritmos para dibujar exquisitos helechos y otras imágenes basadas en fractales, figuras geométricas que se repiten a sí mismas en muchas escalas. Pero era un trabajo que consumía mucho tiempo. Los estudiantes de segundo ciclo "muy espabilados" necesitaban alrededor de 100 horas para cifrar cada figura, recuerda Barnsley.

En 1987, Barnsley y Sloan se pusieron a investigar en técnicas comerciales, basadas en fractales, para comprimir los datos. Fundaron Iterated Systems, adoptaron un helecho como emblema y contrataron un grupo de matemáticos. Entonces, hacia mediados de 1988, "descubrí cierta clase de teoremas que —con ligeras modificaciones— nos permiten compimir automáticamente cualquier imagen", dice Barnsley.

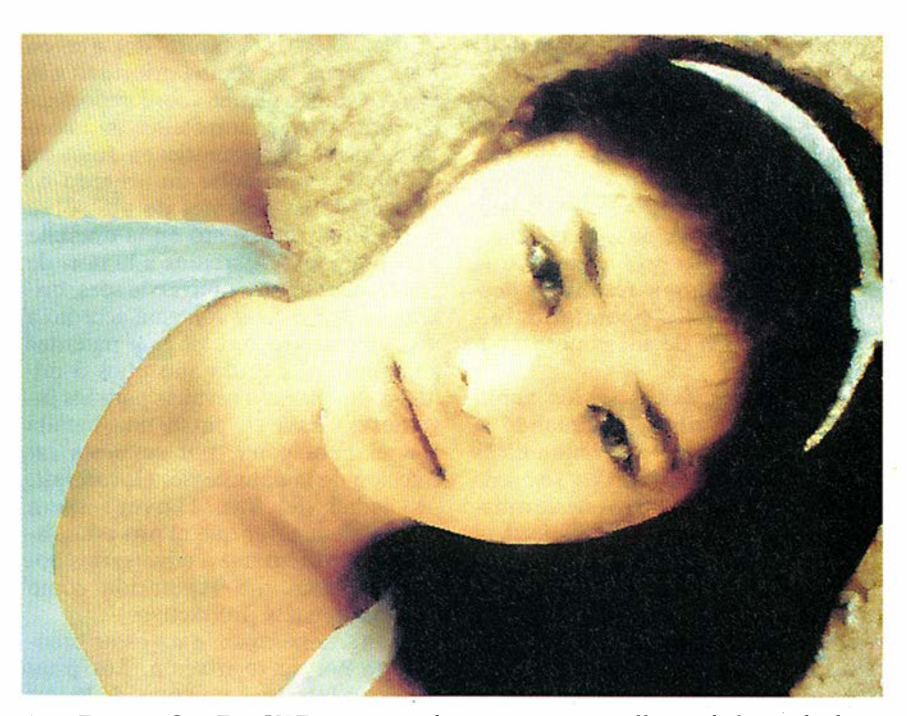
La imagen de la mujer, dice, se cifró en aproximadamente una hora. Los secretos del proceso exacto, sin embargo, son propiedad del autor. Los clientes "tendrán que fiarse de mi reputación", afirma Barnsley. (La compañía ha solicitado dos patentes y está redactando la petición de una tercera.) Su insistencia en el secreto ha irritado a algunos matemáticos, que tienen reservas sobre sus declaraciones. Ponen en duda que sea posible hallar una técnica genérica, basada en fractales, para codificar rápidamente una imagen. El cifrado parece requerir una gran cantidad de métodos matemáticos nuevos.

Antes de que una imagen pueda expresarse con fractales, debe organizarse en un estado de equilibrio. Los investigadores deben hallar y aplicar las reglas para realizar dicho estado. El retrato de la mujer, por ejemplo, llevaba consigo, según Barnsley, unas 1000 transformaciones. Todo lo que se necesita para reconstruir la imagen es una descripción de esas transformaciones; como las reglas se pueden expresar mediante unos pocos parámetros, se pueden obtener relaciones de compresión muy elevadas.

Muchos matemáticos dudan de que todas las imágenes admitan semejante representación, basados en la inexistencia de buenas definiciones y teoremas sobre lo que se puede cifrar eficazmente. La técnica sólo es aprovechable si no requiere un número ingente de transformaciones.

La compañía proyecta expedir sus primeros dispositivos de cifrado en el próximo verano, una vez que Iterated Systems tenga ya a punto la otra mitad del sistema: los descifradores. (Estos, que se conectan a un ordenador personal, descifrarán las imágenes fractales almacenadas en relaciones superiores a 64:1 y desplegarán la imagen en escalas de grises a velocidades de vídeo de 30 cuadros por segundo, adelanta Barnsley.)

Hasta que las cifradoras se expendan en el comercio, los usuarios pueden trabajar con imágenes proporcionadas por Iterated Systems. Los clientes pueden enviar imágenes fijas o vídeos a la compañía, que cifrará las imágenes y devolverá el cifrado vía correo electrónico. Entre los interesados en esta tecnología están los televendedores y los agentes inmobiliarios que deseen mostrar su oferta a



*1. RETRATO DE MUJER constituido por unos tres millones de bytes de datos y comprimido a 5800 bytes mediante una técnica desarrollada por Iterated Systems de Atlanta. Fotografía original de Noriaki Yokosuka/Misako Tanaka.*

través de ordenadores personales. "Estamos en el comienzo de un proceso", dice Sloan, "en el que estamos creando la técnica y levantando el negocio".

## *Empuje conservador*

La penicilina trajo la fama a los dos científicos británicos que la desarrollaron en los años veinte y treinta; gloria aparte, no les dio ni un céntimo. Su gobierno hizo públicos, sin cargo, los detalles del medicamento obtenido de hongos. En el frenesí de la segunda guerra mundial, el departamento de agricultura de los EE.UU. tropezó con el descubrimiento, desarrollando un proceso de fermentación para fabricar penicilina en grandes cantidades. Las firmas británicas se encontraron con que habían de pagar derechos a sus equivalentes del otro lado del Atlántico para emplear el proceso. Escarmentado con la experiencia, el gobierno británico creó una agencia guardiana en 1949 para proteger ideas generadas en sus propias universidades.

Cuatro decenios más tarde, esa oscura rama del servicio civil del Reino Unido, transformada ahora en Grupo de Tecnología Británica (BTG), está preparándose para jugar una importante baza en los mercados que emerjan de la eliminación de viejas barreras políticas y económicas entre Europa occidental y oriental. El BTG confía también en representar a las invenciones norteamericanas en el laberíntico mercado internacional de la propiedad intelectual.

El BTG dice que es la mayor organización en el mundo dedicada exclusivamente a la transferencia de tecnología, es decir, el paso de la técnica del laboratorio al mercado. El personal de BTG, unos 60 especialistas, sigue la pista a los problemas que la industria se afana por resolver y después investiga en los laboratorios británicos en busca de ideas potencialmente útiles.

Los primeros éxitos se los anotó BTG en los años cincuenta, cuando la agencia consiguió patentes mundiales para las cefalosporinas, una segunda generación de antibióticos desarrollados en la Universidad de Oxford, y para el sistema de suspensión de colchón de aire del Hovercraft. Muchos esfuerzos han dado resultados sólo después de años de paciente financiación. El BTG pagó, durante veinte años, la investigación en piretrina, pesticida ambientalmente benigno que hasta 1982 no se convirtió en un éxito de ventas. De manera análoga, la agencia empleó un decenio apoyando el estudio sobre obtención de imágenes por resonancia magnética (IRM), antes de que esta herramienta de diagnóstico médico rindiera apreciables regalías en 1984. En la actualidad, las regalías correspondientes a las patentes de IRM y la piretrina constituyen el 50 por ciento de los ingresos del BTG.

Los fracasos han superado naturalmente el número de éxitos en el siempre arriesgado negocio de invertir en tecnología. Pero, en el camino, BTG cometió otros errores, dicen los observadores. A finales del decenio de 1970, se había atrincherado en la rutina burocrática, abotargada, de difícil trato e indiferente a las preocupaciones del mercado.

Con la llegada de los conservadores al gobierno, se exigieron cotas de rendimiento. La administración Thatcher consideró la posibilidad de privatizarla. Dio también libertad a las universidades y laboratorios nacionales británicos para tratar con otras agencias, a fin de patentar y comercializar su tecnología, acabando así con el confortable monopolio sobre invenciones financiado por el contribuyente británico.

La dama de hierro ha galvanizado el BTG, decidido a demostrar, con 1500 patentes, y regalías anuales de unos 24 millones de libras esterlinas, que está mejor equipado para hacer respetar las patentes que otras organizaciones menores.

Para demostrarlo, Ian A. Harvey, instituido en la jefatura del BTG en 1985, asignó la máxima prioridad a la lucha en cinco pleitos sobre importantes infracciones, entre ellos uno contra la firma norteamericana Johnson & Johnson por violar un acuerdo de licencia de las patentes de BTG sobre IRM. El BTG representó en la querella a 20 patentes, referentes a la IRM, de científicos de tres universidades británicas. Johnson & Johnson accedió a negociar el caso en 1986. Las patentes generaron el pasado año unos 3 millones de libras en regalías para los inventores. Está en curso una batalla con el Pentágono, por supuesta infracción de la patente del Hovercraft, propiedad del BTG. Harvey espera que, al demostrar que el BTG está dispuesto a enfrentarse a adversarios poderosos, adquirirá reputación como litigante tenaz y productivo.

El BTG tiene razón para estar preocupado por su reputación. Sus principales pilares, las patentes de la IRM y la piretrina, expirarán en unos pocos años. De ahí su interés en reforzar su cartera con tecnologías norteamericanas. Consiguió su primer cliente estadounidense importante en marzo de 1989, cuando Johnson & Johnson decidió abandonar el negocio de diagnóstico por imágenes y asignó al BTG la tarea de gestionar más de 100 patentes relacionadas con las imágenes que la compañía poseía en Europa y Extremo Oriente.

## *¿Qué hacer con las monedas del bloque oriental?*

Los asesores de inversión sostienen que la liberalización de Europa oriental constituye la mayor oportunidad para hacer negocios desde los alegres días del Plan Marshall: economías en ruinas y dispuestas a su levantamiento y obreros cualificados impacientes por llenar las fábricas para transformarse también en consumidores. En el proceso se pueden hacer millones de forints, zlotys y koruny (coronas checas). El problema estriba en que nadie quiere forints, zlotys ni koruny.

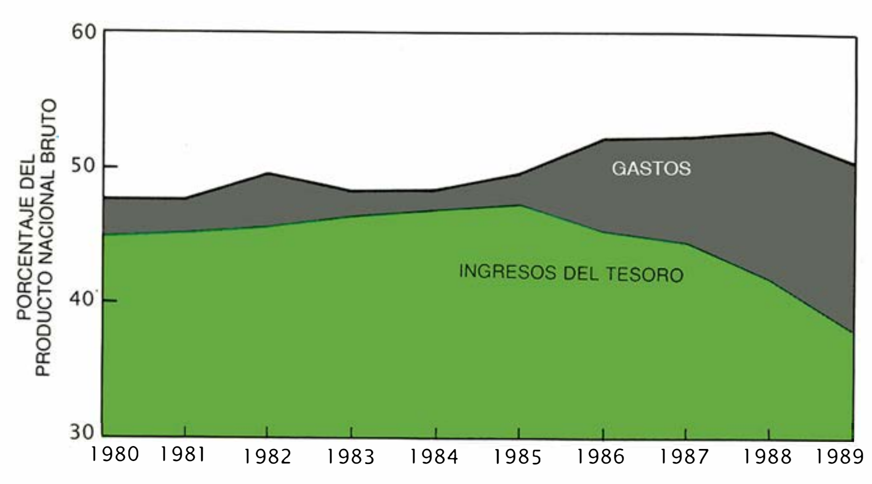
Las monedas convertibles, o "fuertes", se cambian entre sí con libertad y vinculan con sólidos lazos las economías de las respectivas naciones. Las monedas no convertibles (las del bloque oriental y de cierto número de naciones en vías de desarrollo) sólo se cambian con respaldo del gobierno y aíslan sus economías locales de las del resto del mundo.

Las monedas no convertibles desalientan también la inversión extranjera. Si una empresa occidental gasta marcos alemanes, con los que puede comprar cualquier cosa, en construir una fábrica en Rumanía, recibe los beneficios en lei, que puede gastar únicamente en Rumanía. A menos que la compañía obtenga permiso para volver a convertir los lei en marcos, el resultado es como si hubiese perdido ese dinero. Pocas empresas aceptan tales términos; por ello, los gobiernos que tienen monedas no convertibles establecen tipos de cambio artificialmente altos para sonsacar tanta moneda convertible como puedan a quienes han de realizar negocios allí.

Pero ahora los países del bloque oriental quieren atraer la inversión extranjera; necesitan el dinero para reconstruir sus maltrechas economías y proporcionar a sus ciudadanos un nivel de vida comparable con el de occidente. Lo primero que deberían hacer, en opinión de algunos economistas, es tornar convertibles sus monedas. Uno de los principales defensores de este punto de vista es Jeffrey D. Sachs, "lumbrera de Harvard",

que antes logró mejoras en la economía de Bolivia y ha emprendido ahora la restructuración de la de Polonia.

Otros economistas sostienen que hacer, en este momento, convertibles las monedas del bloque oriental es prematuro, si no peligroso. (Polonia puede ser la excepción, dice Richard E. Quandt, de la Universidad de Princeton, por lo deteriorada que ya está su economía.) Antes de hacer sus monedas convertibles, dice Edward A. Hewett, de la Institución Brookings de Washington, las naciones del bloque oriental deben abordar el tránsito de un sistema económico protegido, en el que el dinero no importa, a la economía mundial, en la cual sí importa. En el Este, el dinero ha servido de medio pasivo de contabilidad para seguir la pista a las transacciones, no como un procedimiento de cambio por medio del cual los precios de las mercancías se ajustan a su valor real.




*2. EL DEFICIT PRESUPUESTARIO SOVIETICO ha aumentado bruscamente desde el comienzo de la* perestroika. *Según Ronald I. McKinnon, de la Universidad de Stanford, la caída de los ingresos (*verde*) es una consecuencia natural de la liberalización: los "beneficios" de las empresas previamente controladas por el estado ya no afluyen a Moscú. Fuente: Planecon.*

La Unión Soviética, por ejemplo, explica Ronald I. McKinnon, de la Universidad de Stanford, tiene al menos tres clases de "moneda", ninguna de ellas convertible en las otras. En primer lugar, los individuos pueden ahorrar dinero en el papel o moneda expedidos por el gobierno; no así las empresas. En segundo lugar, los beneficios que las empresas generan van a parar a cuentas en el banco estatal; los distintos superávits se destinan a objetivos diversos, el gastar cualquiera de ellos requiere la aprobación del estado. Finalmente, el banco estatal presta a las empresas moneda "blanda" con la que pueden adquirir materiales para cumplir sus cuotas de producción. Cuando el gobierno asigna todas las mercacías por decreto, el sistema funciona bien, pero en un mercado libre sus efectos son "perversos", razona McKinnon. Las empresas que pierden dinero pueden siempre obtener mas préstamos, y las que producen beneficios no pueden invertir lo que ganan para aumentar su rendimiento.

Otra medida crucial es la de crear un sistema de impuestos adecuado. La liberalización de las economías socialistas, explica McKinnon, ha conducido a enormes déficits y al caos económico porque los comunistas reformadores no sabían qué tenían entre manos.

Cuando el estado es propietario de casi todo, el dinero de los sectores "beneficiosos" —los que hacen cosas que la gente compra— va a parar a financiar los deficitarios: la defensa nacional. Un sistema explícito de impuestos resulta innecesario, por la sencilla razón de que el gobierno puede establecer los salarios y los precios en niveles que le dejen el margen que necesita. En ausencia de impuestos explícitos, la transición a la empresa privada representa una caída vertical de los ingresos. En la Unión Soviética, los ingresos gubernamentales cayeron del 47 por ciento del producto nacional bruto en 1985 al 38 por ciento en 1989; el déficit presupuestario aumentó casi en la misma cuantía, de 2,7 por ciento del producto nacional bruto a 13,1 por ciento.

Sin sistemas monetarios e impositivos adecuados, liberalizar la moneda del bloque oriental provocaría un cuadro económico más sombrío de lo que es hoy en esos países. En un mercado libre, el valor de las monedas orientales caería, elevando el precio de los bienes importados y desencadenando potencialmente la hiperinflación. Los inversores extranjeros encontrarían tan barato adquirir activos locales que dejarían fuera de juego a los empresarios del país.

Pero semejante situación coloca a los países del Este ante un dilema. Si la liberación de la conversión del forint o la koruna empantana la economía local, ¿de dónde sacar el dinero necesario para la inversión? Algunos economistas, como McKinnon y Janos Hórvath (fugaz presidente del Consejo Nacional de Reconstrucción Económica de Hungría en 1956) sugieren fuentes internas: las empresas deberían abandonar el paraguas de los créditos y verse obligadas a financiar su expansión con los beneficios obtenidos. Si esta draconiana solución opera con los buenos resultados conseguidos en Alemania occidental en las postrimerías de los años cuarenta, las compañías elevarán los precios y los consumidores financiarán la inversión sacando ahorros que de otra manera habrían permanecido ociosos. Una vez la economía florezca, habrá tiempo para hablar de monedas convertibles.

De nuevo entonces, podría llegar la hora de la inversión extranjera en la reconstrucción del bloque oriental, como ocurrió en Europa bajo el Plan Marshall. Si es así, dice Hewett, medidas próximas a la convertibilidad permitirán todavía a los occidentales contemplar sus beneficios en moneda fuerte. Algunas compañías han practicado el intercambio durante años; probablemente el mejor ejemplo es el lucrativo comercio de bebidas refrescantes por vodka. Otra alternativa podría ser las subastas de moneda fuerte. Las empresas podrían convertir regularmente el dinero local en yens o marcos, pero no de la manera fluida y dúctil que caracteriza la compraventa de moneda en Occidente.

Existiendo tales alternativas, obstinarse en monedas convertibles podría resultar nocivo. Hay otras barreras contra la inversión; están pendientes de resolución los contenciosos sobre derechos de propiedad, por ejemplo. Una vez que tales problemas se hayan resuelto, Europa oriental podría encerrar un atractivo no menor que el de otras zonas en vías de desarrollo; sus carreteras e infraestructura son mejores, sus trabajadores tienen mejor formación y sus salarios no serán prohibitivos.

# Cómo producen oxígeno las plantas

*El mecanismo bioquímico denominado reloj oxidante del agua confiere a las plantas y a ciertas bacterias la capacidad de utilizar la energía solar para romper las moléculas de agua en oxígeno molecular, protones y electrones*

Govindjee y William J. Coleman

Puesto que el oxígeno molecular es indispensable para los seres humanos es fácil olvidar que los organismos más sencillos vivieron sin él durante cientos de millones de años. Para los primitivos organismos anaeróbicos, el oxígeno era una sustancia tóxica que podía sustraer electrones esenciales de los componentes moleculares de sus células. Incluso puede parecer sorprendente que muchas de aquellas células anaeróbicas desarrollaran cierto tipo de fotosíntesis en ausencia de oxígeno, ya que ésta es el proceso que produce todo el oxígeno de la atmósfera. El mecanismo preciso por el que se genera oxígeno en la fotosíntesis estuvo envuelto en el misterio, aunque ahora lo conocemos ya con bastante detalle. Se trata de un "reloj oxidante del agua" que genera una molécula de oxígeno cada cuatro pulsaciones.

Sin embargo, el objetivo principal de la fotosíntesis es hacer posible que las células conviertan el dióxido de carbono en carbohidratos con la energía solar que absorben. El hecho de que la producción de oxígeno en sí no sea esencial explica por qué las células anaeróbicas pudieron realizar fotosíntesis en tiempos remotos sin desprender oxígeno molecular y por qué lo han seguido haciendo así hasta hoy.

Pero si el oxígeno es tóxico, ¿por qué y cómo empezaron las plantas verdes y sus ancestros a producirlo mediante la fotosíntesis? La respuesta a la primera pregunta radica en el metabolismo energético. La vida en la Tierra depende de la energía que suministra la luz solar, pero las células no pueden emplear o almacenar la energía luminosa directamente, sino que tienen que convertirla en energía química, más fácil de utilizar. Los electrones son "moneda" de uso común en la conversión biológica de la energía: muchas de las reacciones bioenergéticas de las células pueden explicarse en términos de transferencia de electrones entre moléculas.

Por tanto, las células necesitan para vivir una fuente de electrones. Las bacterias fotosintéticas anoxigénicas oxidan, es decir, extraen electrones de ácidos orgánicos y simples compuestos inorgánicos. Estas sustancias son, sin embargo, relativamente infrecuentes. En consecuencia, las bacterias anoxigénicas que han sobrevivido hasta hoy se desarrollan solamente en fuentes sulfúreas, fondos lacustres y ambientes similares, en los que abundan estos tipos de moléculas.

Hace unos tres mil millones de años, algunas células fotosintéticas aprendieron a prosperar prácticamente en cualquier tipo de entorno natural, sacando electrones de una sustancia casi ubicua: el agua ($H_2O$). Estas células desarrollaron la facultad de disociar pares de moléculas de agua en electrones, protones ($H^+$, o núcleos de hidrógeno) y oxígeno molecular ($O_2$). Los electrones y protones resultaban energéticamente útiles; el oxígeno era simplemente un producto de desecho. En resumen, el desprendimiento de oxígeno significó un hito para los organismos fotosintéticos, y no porque el $O_2$ fuera importante en sí mismo, sino porque las células fotosintéticas pudieron utilizar el agua e invadir ambientes más diversos.

Mucho más compleja es la pregunta sobre cómo las células producen oxígeno. El desarrollo de la capacidad de utilización del agua como fuente de electrones no fue cosa fácil y exigió varias modificaciones en el mecanismo fotosintético preexistente. Puesto que las moléculas de agua sólo ceden electrones a regañadientes, el oxidante —o molécula aceptora de electrones— relativamente débil que las bacterias fotosintéticas anoxigénicas eran capaces de generar a expensas de la luz solar había de ser remplazado por otro mucho más potente. Pero incluso así, la energía de un fotón, o cuanto de luz, visible no es suficiente para romper una molécula de agua. El problema pudo resolverse aprovechando la energía de cuatro fotones para escindir dos moléculas de agua, liberándose en el proceso cuatro protones y cuatro electrones. Este mecanismo creó, sin embargo, otra dificultad, ya que el

GOVINDJEE y WILLIAM J. COLEMAN han colaborado en estudios sobre el mecanismo de la producción de oxígeno por las plantas verdes. Govindjee nació en Allahabad, India; recibió su grado de doctor por la Universidad de Urbana en 1960, actualmente es profesor de biofísica y biología vegetal en la Universidad de Illinois en Urbana-Champaign y ha escrito muchos trabajos, revisiones y libros sobre fotosíntesis. El doctor Coleman es becario de la Fundación Nacional de la Ciencia e investiga sobre fotosíntesis bacteriana en el departamento de química del Instituto de Tecnología de Massachusetts.

**1. BURBUJAS DE OXIGENO en las hojas de una planta verde sumergida; es una señal de que se está desarrollando la fotosíntesis. El oxígeno ($O_2$) es un producto de la reacción promovida por la luz en la que se eliminan cuatro electrones ($e^-$) y cuatro protones ($H^+$) de cada par de moléculas de agua. Las bacterias fotosintéticas anoxigénicas, incapaces de escindir así las moléculas, obtienen los electrones de otras fuentes.**

aparato fotoquímico sólo puede manipular los electrones de uno en uno.

Las células fotosintéticas resolvieron dicho problema al desarrollar un catalizador especial para romper el agua, el denominado reloj oxidante del agua: un mecanismo bioquímico único de "trinquete" para estabilizar los estadios intermedios de la reacción de lisis del agua, de manera que los electrones pudieran transferirse de uno en uno. En los últimos años se ha avanzado mucho en el conocimiento del reloj oxidante del agua y de su intervención en el proceso fotosintético global.

En las plantas superiores, las reacciones primarias de la fotosíntesis tienen lugar en las membranas tilacoidales especializadas, incluidas en el interior de los cloroplastos, unas estructuras celulares. Embebidos en dichas membranas tilacoidales se encuentran varios complejos proteicos, cada uno de los cuales contribuye a la reacción fotosintética global. La formación de $O_2$ acontece totalmente dentro del complejo de proteínas y pigmentos que se conoce como fotosistema II, propio de las células de todos los organismos fotosintéticos oxigénicos: cianobacterias, algas y plantas con clorofila.

La función principal del fotosistema II es la de actuar como un pequeño condensador que almacena la energía mediante la separación y la estabilización de cargas positivas y negativas a uno y otro lado de la membrana tilacoidal. El proceso lo realiza un conjunto de pigmentos especializados del fotosistema II, que absorbe un fotón y convierte eficientemente su energía luminosa en una amplia separación de cargas.

La sincronización de movimientos en el complejo proceso de conversión de la energía luminosa en una separación de cargas requiere la colaboración de muchos polipéptidos y proteínas especializados del fotosistema. Los polipéptidos son polímeros lineales de muchos residuos de aminoácidos, a menudo de cientos de ellos, dispuestos en una secuencia determinada. Las proteínas están constituidas por uno o más polipéptidos plegados en intrincadas y ordenadas estructuras.

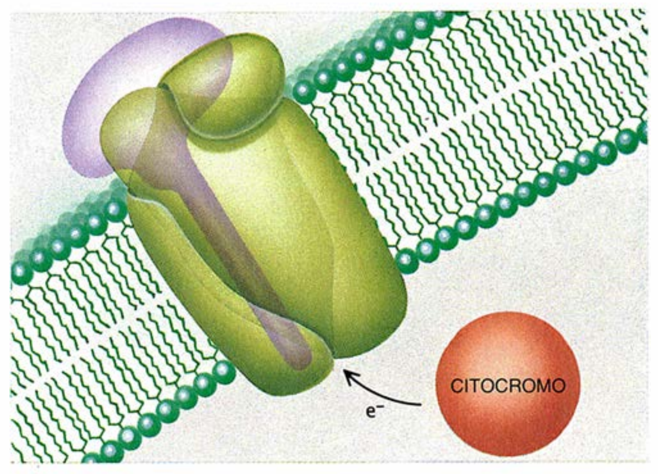
Las reacciones de transferencia de electrones en el fotosistema II se desarrollan en el centro de reacción. Los principales componentes estructurales del centro de reacción son dos polipéptidos de gran tamaño, llamados D1 y D2, y una proteína menor, conocida por citocromo $b_{559}$. Unidos a la cara interna de la membrana del tilacoide encontramos un polipéptido de 33 kilodalton de peso molecular y al menos otros dos de diferentes pe-



sos. Estos polipéptidos sirven como matriz estabilizadora de los pigmentos y otras moléculas del fotosistema II que llevan a cabo las reacciones de transferencia de electrones y producción de oxígeno. Otros polipéptidos de pequeño tamaño se asocian igualmente con el fotosistema II, pero sus funciones se desconocen todavía. Varios iones orgánicos y átomos cargados —manganeso, cloruro, calcio, hierro y bicarbonato— están implicados en la catálisis de la transferencia de electrones, en el mantenimiento de la estructura proteica o en la regulación de la actividad del fotosistema.

Además, un gran número de moléculas de clorofila "antena" captan la energía luminosa y la canalizan eficientemente hasta el centro de reacción, con el que se asocian varios cientos de moléculas de pigmentos antena.

Al ser tan complicada la estructura del fotosistema II, la mayor parte de lo que hoy se conoce sobre el mismo




**2. FOTOSISTEMA II: complejo de pigmentos y proteínas de la membrana tilacoidal de los cloroplastos que capta la energía luminosa y produce oxígeno. El flujo de electrones a través del fotosistema (*arriba*) se promueve con la absorción de un fotón por un par especial de clorofilas (P680). El proceso implica muchas otras moléculas, como la feofitina (Pheo) y dos tipos de quinonas ($Q_A$, $Q_B$), así como átomos metálicos e iones, tales como manganeso (Mn), cloruro ($Cl^-$) y calcio ($Ca^{2+}$). En este modelo simplificado se han omitido varias moléculas de pigmento en aras de la claridad. Las reacciones del fotosistema II en el lado interior de la membrana extraen cuatro electrones de dos moléculas de agua y liberan una molécula de oxígeno; estas reacciones constituyen el reloj oxidante del agua. Los protones liberados en estas reacciones contribuyen a la síntesis de adenosín trifosfato. El fotosistema de las bacterias anoxigénicas (*derecha*), más simple, carece del reloj oxidante del agua; otros compuestos distintos del agua suministran electrones al fotosistema mediante una proteína citocrómica.**

procede, de hecho, de los estudios realizados sobre los correspondientes complejos de las bacterias fotosintéticas. El trabajo de Johann Deisenhofer, Robert Huber y Hartmut Michel, quienes determinaron la estructura del centro fotosintético de reacción de la bacteria *Rhodopseudomonas viridis*, les mereció el premio Nobel de química de 1988.

Hay muchas diferencias entre los complejos fotosintéticos de las bacterias y de las plantas verdes. Como se ha mencionado antes, las bacterias no producen oxígeno molecular; además, estos microorganismos no utilizan clorofila sino bacterioclorofila, pigmento que absorbe principalmente luz de longitudes de onda mucho más largas y que es un oxidante mucho más débil. Sin embargo, los complejos fotosintéticos bacterianos catalizan la reacción clave que convierte la luz en un gradiente electroquímico de protones a través de la membrana biológica [véase "Mecanismo molecular de la fotosíntesis", por Douglas C. Youvan y Barry L. Marrs; INVESTIGACIÓN Y CIENCIA, agosto de 1987].

Sobre la base de estos estudios con bacterias se puede decir que el mecanismo de transporte de electrones del centro de reacción del fotosistema II comprende cinco componentes: el pigmento clorofila, que actúa como donador primario de un electrón; un donador secundario de electrones denominado Z, que reduce la clorofila, es decir, que restituye el electrón que pierde la clorofila; la feofitina, el pigmento que acepta un electrón de la clorofila; la plastoquinona $Q_A$, un aceptor primario de electrones, y la plastoquinona $Q_B$, un aceptor secundario de electrones.

Se cree que el pigmento clorofila del centro de reacción está constituido por un par especial de moléculas de clorofila que, aunque parecen ser químicamente idénticas a muchos de los pigmentos antena, difieren en su función. Dicho pigmento se denomina P680 porque absorbe con más intensidad la luz de 680 nanómetros de longitud de onda.

En 1988, el trabajo de Bridgette A. Barry y de Richard J. Debus, de la Universidad estatal de Michigan, y de Willem F. Vermaas y sus colaboradores, de la de Arizona, contribuyó a identificar el factor Z como tirosina, uno de los aminoácidos del polipéptido D1. La quinona $Q_A$ se encuentra fuertemente unida al complejo del fotosistema II, mientras que la quinona $Q_B$ puede difundir libremente entre los complejos proteicos de la membrana después de aceptar dos electrones (o, en términos químicos, de ser doblemente reducida).

Durante la fotosíntesis, los pigmentos antena absorben un fotón y canalizan su energía hacia el P680 del centro de reacción, en el que la energía de excitación se traduce en una separación de cargas cuando P680 alcanza un estado excitado y transfiere rápidamente un electrón a una molécula vecina de feofitina [*véase la figura 3*]. La feofitina posee, entonces, una carga negativa en exceso, mientras que hay un "agujero" con carga positiva en P680 debido a la pérdida del electrón: P680 queda como $P680^+$. La separación de cargas se hace mayor cuando la feofitina cede su electrón extra a $Q_A$ y la distancia aumenta aún más cuando Z dona un electrón a $P680^+$, quedándose con carga positiva, y $Q_A$ transfiere su electrón adicional a $Q_B$.

La transferencia de carga opera con rapidez, especialmente la transferencia inicial de un electrón desde el estado excitado de P680 hasta la feofitina, lo que ocurre en una pocas billonésimas de segundo. Esto fue demostrado por uno de nosotros (Govindjee) en colaboración con Michael R. Wasielewski y Douglas G. Johnson, del Laboratorio Nacional de Argonne, y Michael Seibert, del Instituto de Investigación de Energía Solar de Golden, Colorado.

La transferencia secuencial de electrones alcanza su objetivo al conseguir distanciar lo más posible las cargas positivas y negativas, que se atraen mutuamente. No obstante, el ciclo fotosintético del fotosistema II no se completa hasta que todos los componentes del mecanismo de reacción recuperan su estado neutro inicial y se encuentran en disposición de empezar, una vez más, el proceso de separación de cargas. ¿Cómo pierde $Q_B$ su carga negativa y cómo recupera Z el electrón perdido?

En el extremo del sistema en el que se encuentra $Q_B$, la respuesta es relativamente sencilla. Después de que $Q_B$ ha atrapado dos electrones y dos protones mediante dos ciclos de absorción de fotones, la molécula de $Q_B$ doblemente reducida se difunde hacia fuera del complejo del fotosistema II y es sustituida por otra molécula de $Q_B$ oxidada. Los electrones y protones asociados a la molécula de $Q_B$, que posee absoluta libertad de movimiento, son transportados a otro complejo de la cadena fotosintética. Los protones liberados en la cara interna de la membrana tilacoidal terminarán por aprovecharse en la síntesis de adenosín-trifosfato, un compuesto de almacenamiento de energía esencial para el metabolismo celular.

En el extremo opuesto del fotosistema II resulta mucho más difícil que Z obtenga el electrón necesario para recuperar su estado inicial. El electrón debe proceder de alguna sustancia oxidable que esté disponible en el entorno de la célula. Los ácidos orgánicos (tales como acetato, malato y succinato) y los compuestos inorgánicos sencillos (tales como sulfuro y tiosulfato) pueden ser buenas fuentes de electrones y, de hecho, son las que utilizan las bacterias fotosintéticas anoxigénicas; en estas bacterias, que carecen de compuesto Z, es una proteína del citocromo el donador que suministra electrones al par especial de clorofilas oxidadas del centro de reacción.

Una molécula mucho más abundante que los ácidos orgánicos —y, por tanto, una fuente de electrones potencialmente más rica— es el agua. Sin embargo, aunque la capacidad de oxidación de $P680^+$ es alta, no basta por sí sola para despojar de sus electrones a la molécula de agua.

El problema es que la reacción de oxidación del agua libera cuatro electrones simultáneamente, mientras que $P680^+$ sólo puede aceptar electrones de uno en uno. En consecuencia, los científicos vieron con claridad hace unas décadas que debía existir un sitio catalítico cerca de Z y P680 que pudiera prolongar así la reacción de oxidación. Este catalizador de la lisis del agua ha de ser capaz de asociar pares de moléculas de agua y estabilizarlas durante un proceso gradual de oxidación en el que los electrones son transferidos de uno en uno. Las investigaciones sobre este mecanismo condujeron al descubrimiento del reloj oxidante del agua.

La observación de que no todos los electrones alcanzan la molécula de clorofila de $P680^+$ a la misma velocidad resultó decisiva para la comprensión del funcionamiento de este mecanismo. En realidad, el tiempo observado de transferencia de electrones varía de manera periódica; esto se demostró mediante experimentos en los que las membranas con los centros de reacción del fotosistema II se mantienen en la oscuridad y se exponen a continuación a breves pulsos de luz. Cada destello no sólo es de gran intensidad sino también lo más breve posible, de tal forma que introduce un solo fotón, por término

medio, en el fotosistema. El resultado observado es que $P680^+$ recupera el electrón en la oscuridad a velocidades diferentes, dependiendo del número de pulsos de luz.

Por ejemplo, el tiempo necesario para que la mitad de moléculas de $P680^+$ revierta a P680 es de unos 20 nanosegundos después del primer y quinto destellos, pero es muy superior después del segundo, tercero y cuarto. El cambio en el tiempo de recuperación varía cíclicamente cada cuatro pulsos. Esta periodicidad sugiere la existencia de una reacción cíclica de cuatro etapas para donar electrones al centro de reacción.

Tales estudios sobre el comportamiento de P680 revistieron especial interés. Pierre Joliot, del Instituto de Biología Fisicoquímica de París, había demostrado con anterioridad, en 1969, que había una periodicidad de cuatro en la producción fotosintética de oxígeno. Utilizando un aparato con un electrodo de platino altamente sensible, que respondía a cantidades mínimas de $O_2$, Joliot midió la cantidad de oxígeno que se desprendía después de una serie de pulsos de luz. Observó que no había desprendimiento de $O_2$ después del primer pulso, y nada (o muy poco) después del segundo, pero obtuvo una cantidad máxima tras el tercero. A continuación, la amplitud de la señal de $O_2$ oscilaba con una periodicidad de cuatro hasta que las diferencias disminuían gradualmente.

Bessel Kok, de los laboratorios Martin Marietta en Baltimore, sugirió en 1970 una hipótesis sencilla para explicar los resultados de Joliot; propuesta que se conoce como el reloj, o ciclo, de oxidación del agua. Kok propuso que el complejo productor de oxígeno en el fotosistema II podría existir en varios estados transitorios de oxidación diferentes: estados S. Aunque no pudo definir la naturaleza química de los estados S con precisión, sugirió que cada estado S participaba una sola vez en un mecanismo cíclico de cuatro etapas.

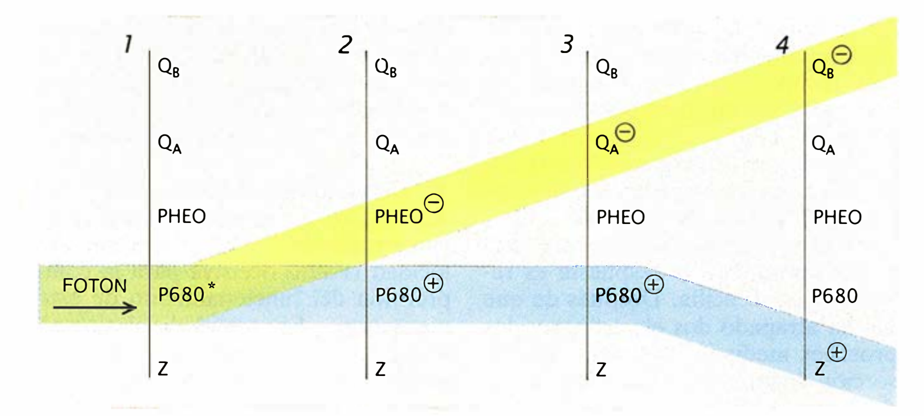
Kok predijo que, en la oscuridad, el reloj se hallaba en uno de los dos estados S iniciales: $S_0$, o bien $S_1$. El estado predominante (y más estable) es $S_1$, que tiene un equivalente de oxidación más que $S_0$; en otras palabras, el complejo de moléculas correspondiente a $S_1$ tiene un electrón menos que el complejo $S_0$. Se desconoce la base química para el predominio de $S_1$.

Después de un pulso luminoso, P680 pasa a $P680^+$. Debe terminar en una fase reducida con un electrón. Kok supuso que el reloj debía sufrir algún cambio que lo impulsaba al próximo, y más alto, estado de oxidación: el reloj que comienza en $S_1$ pasa a $S_2$, y el que empieza en $S_0$ pasa a $S_1$. La transición tiene lugar al liberarse un electrón del reloj y convertir $P680^+$ en P680. Un segundo destello crea otro $P680^+$ e impulsa el estado $S_2$ hacia el $S_3$, y un tercer pulso de luz convierte el estado $S_3$ en $S_4$. Al alcanzar el estado $S_4$, el reloj ha liberado cuatro electrones, encontrándose dispuesto para completar la reacción de lisis del agua. El reloj elimina entonces cuatro electrones de las dos moléculas de agua que tenía asociadas, desprende $O_2$ y vuelve desde $S_4$ a $S_0$, posibilitando la entrada en un nuevo ciclo.

Esta situación puede compararse a la de un jugador de béisbol: el jugador debe marcar secuencialmente las cuatro "bases" hasta acabar donde empezó. Si el jugador falla una marca puede verse forzado a retirarse; de forma similar, existe la posibilidad de que el reloj no progrese regularmente de un estado al siguiente. Puede ocurrir en unos pocos casos, por ejemplo, que $S_1$ no cambie a $S_2$ después de un pulso luminoso debido a que el fotosistema no aprovechó bien el fotón. También es baja la probabilidad de que un fotosistema absorba dos fotones durante un destello (si los pulsos no son suficientemente cortos) y el reloj oxidante del agua avance en un solo paso de $S_1$ a $S_3$ vía $S_2$.

El mecanismo de Kok explicó las observaciones de Joliot sobre el comportamiento del reloj productor de oxígeno. Debido a que la mayoría de los relojes en una muestra adaptada a la oscuridad se encuentra en el estado $S_1$, el desprendimiento máximo de $O_2$ tiene lugar después del tercer pulso de luz, cuando el reloj cambia de $S_3$ a $S_4$ y a $S_0$, y libera espontáneamente oxígeno. El reloj que empezó en el estado $S_0$ libera $O_2$ tras el cuarto destello, lo que explica que entonces se dé una pequeña liberación de $O_2$.

Estos "errores", que ocurren al azar cuando unos pocos relojes fallan en su avance durante un pulso de luz, o cuando avanzan dos estados S, explican la disminución gradual de las oscilaciones en el desprendimiento de $O_2$. Estos procesos hacen perder lentamente la sincronía de los ciclos de los relojes de la muestra. Después de muchos pulsos de luz hay un equilibrio tal que el número de relojes en los estados $S_0$, $S_1$, $S_2$ y $S_3$ llega a ser aproximadamente igual, y la producción de oxígeno después de cada destello permanece estable. La situación es análoga a la de una habitación llena de relojes de péndulo: de partida, todos pueden dar sincrónicamente una fuerte campanada a la hora exacta, pero, a medida que los relojes adelantan o atrasan, la habitación empieza a resonar con un suave y continuo carillón.



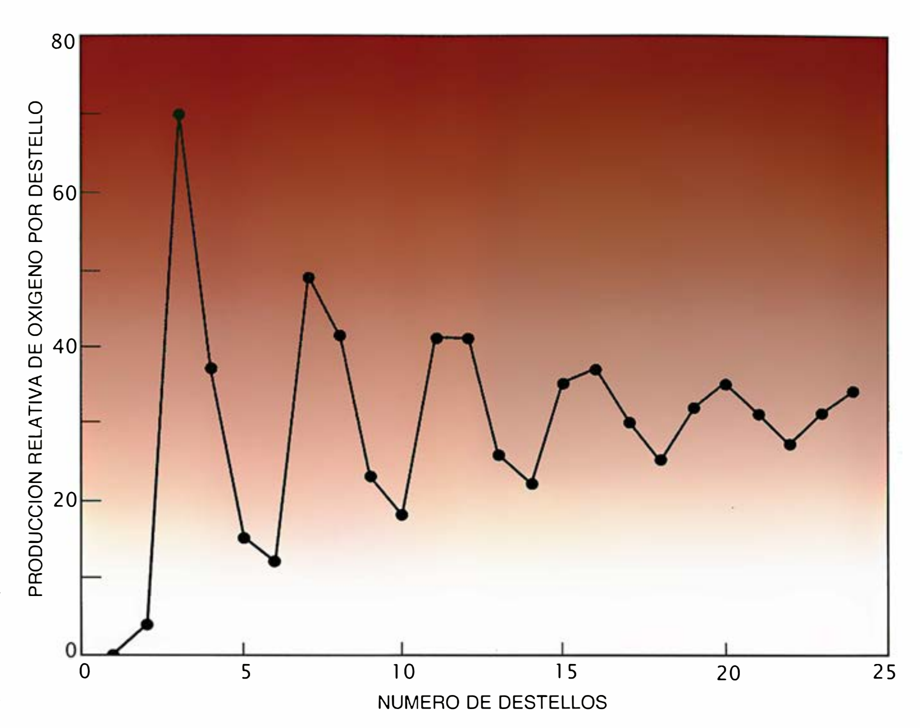

**3. TRANSFERENCIA SECUENCIAL DE ELECTRONES del centro de reacción del fotosistema II; dicho proceso almacena parte de la energía luminosa en forma de cargas positivas y negativas separadas. Tras la absorción de un fotón, el pigmento P680 alcanza un estado excitado y pasa a P680* (*1*). El P680* cede un electrón a la molécula de feofitina (Pheo) y se queda con un electrón menos, o sea, con una carga positiva (*2*). La carga negativa de la feofitina pasa a una molécula de quinona primaria, $Q_A$ (*3*). Finalmente, el $P680^+$ acepta un electrón de Z, un aminoácido del polipéptido D1, y $Q_A$ transfiere su electrón adicional a $Q_B$, otra molécula de quinona (*4*). La cadena de transferencia de electrones vuelve a su estado original cuando Z recibe un electrón del reloj de oxidación del agua, y la molécula de $Q_B$ doblemente reducida es sustituida por otra molécula de quinona oxidada.**

El descubrimiento, de Joliot y Kok, del reloj oxidante de agua vino a sustituir la caja negra de producción de oxígeno por un nuevo mecanismo teórico. La teoría no explicaba, sin embargo, la estructura física del reloj, ni las interacciones de éste con las moléculas de agua. A partir de enton-

ces empezó una larga búsqueda de la naturaleza química del acumulador de cargas en el reloj: el material o materiales cuyos estados de oxidación variable dan lugar a cada uno de los estados S.

Desde el principio se supuso que esta entidad química tan esquiva era un átomo metálico. Los átomos de metales de transición unidos a proteínas, tales como el manganeso, el hierro y el cobre, son buenos candidatos para catalizar las reacciones de oxidación-reducción, gracias a su capacidad para dar y aceptar electrones alternativamente.

Se cree que el manganeso (Mn) forma parte de la estructura del acumulador de cargas ya que, como se sabe desde hace tiempo, la producción de $O_2$ no tiene lugar a no ser que haya cuatro átomos de manganeso en el fotosistema II por cada molécula de P680. Se sabe que el manganeso cataliza las reacciones de transferencia de electrones en otras enzimas. Además, puede adquirir varios estados de oxidación relativamente estables, desde +2 a +7; es decir, los iones de manganeso pueden compartir entre dos y siete electrones con otros átomos. Cuando el metal se une a una molécula de gran tamaño, como una proteína, estos estados de oxidación suelen abreviarse como Mn(II), Mn(III) y así sucesivamente.




**4. LA PRODUCCION DE OXIGENO POR MEMBRANAS FOTOSINTETICAS expuestas a una serie de destellos oscila con una periodicidad de cuatro, siendo máxima tras el tercer destello y de nuevo cuatro destellos después, pero la amplitud va gradualmente disminuyendo conforme aumenta el número de destellos. La aparición de los picos y el amortiguamiento de las oscilaciones del gráfico superior se explican mediante el ciclo de cuatro etapas del denominado reloj oxidante del agua.**

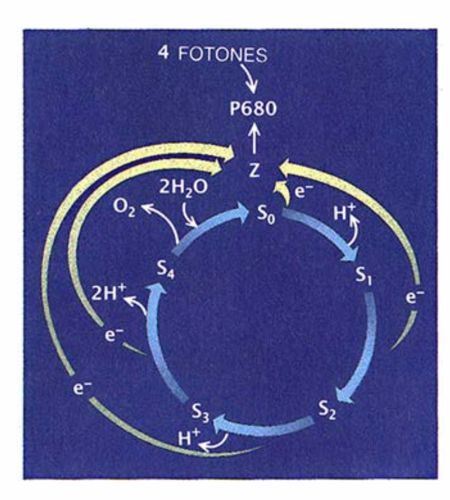
Las proteínas que contienen metales se han analizado en profundidad mediante diversas técnicas espectroscópicas, denominadas así porque algunos complejos metálicos absorben determinados tipos de radiación electromagnética. Si la absorción se mide cuidadosamente, la determinación puede servir como una "huella dactilar" espectroscópica del metal asociado a la proteína y proporcionar pistas acerca de su estructura nuclear o electrónica. La espectroscopía resulta muy adecuada para el estudio de los compuestos de manganeso. Muchos de los complejos de manganeso de interés biológico son "paramagnéticos": el átomo de manganeso contiene electrones con espines desapareados, y estos electrones, como si fueran pequeñas barritas magnéticas, o imanes, interaccionan de manera intensa con los campos magnéticos externos aplicados.

Varias técnicas de medición altamente sensibles, en particular la espectroscopía de resonancia paramagnética del electrón (RPE), han aprovechado las propiedades paramagnéticas del manganeso. Mediante RPE se han estudiado los cambios en la configuración electrónica del complejo de manganeso que siguen a la absorción de luz por el fotosistema II. Otra técnica que ha proporcionado información ha sido la espectroscopía de resonancia magnética nuclear (RMN), que permite medir de forma indirecta las propiedades de los átomos de manganeso registrando los protones de las moléculas de agua que están en contacto con ellos. A mediados de los años setenta, Thomas J. Wydrzynski, entonces en la Universidad de Illinois en Urbana-Champaign, introdujo la utilización del RMN en el estudio de los cambios dinámicos que ocurren en el estado de oxidación del manganeso.

Las técnicas de espectroscopía de rayos X han contribuido en buena medida al estudio de los estados de oxidación y del entorno físico de los átomos de manganeso en el fotosistema II. Otras investigaciones sobre la composición química de los estados S han empleado técnicas de espectroscopía óptica, ya que los complejos de manganeso tienen bandas específicas de absorción en la región ultravioleta del espectro electromagnético.

Es necesario señalar, sin embargo, que, a pesar de la gran diversidad de técnicas espectroscópicas disponibles, dos importantes obstáculos dificultan considerablemente el estudio de la membrana fotosintética. Primero, la complejidad de la membrana, muchos de cuyos componentes tienen espectros de absorción que se solapan. Segundo, los datos obtenidos sobre el reloj oxidante del agua mediante análisis espectroscópicos no pueden interpretarse de forma concluyente, ya que ni la estructura ni la naturaleza química del complejo del fotosistema II se conocen con exactitud. En consecuencia, no se comprende todavía de manera tajante la naturaleza química de los distintos estados S, si bien ha sido posible desarrollar un modelo provisional.

Está claro que los átomos de manganeso sufren cambios dinámicos, verbigracia, cambios en sus estados de oxidación, durante las transiciones entre los estados S. Se ha observado una periodicidad de cuatro en los cambios de los estados de oxidación del manganeso, como había sugerido el modelo de Kok. Un hallazgo sorprendente fue que los átomos de manganeso no se oxidan de forma progresiva a lo largo del ciclo. Mientras que el manganeso está más oxidado en $S_2$ que en $S_1$, y en $S_1$ más que en $S_0$, no hay cambio apreciable en su estado de oxidación entre $S_2$ y $S_3$. Parece, por tanto, que la carga positiva que adquiere el reloj durante su transición de $S_2$ a $S_3$ debe acumularse




**5. RELOJ OXIDANTE DEL AGUA, mecanismo cíclico que proporciona electrones a las clorofilas P680 del centro de reacción del fotosistema II. Cada vez que el pigmento P680 absorbe un fotón, el reloj avanza un estado S, o estado de oxidación, y, por tanto, libera un electrón ($e^-$). Cuando el reloj alcanza el estado $S_4$, libera espontáneamente una molécula de oxígeno ($O_2$) y vuelve al estado $S_0$, cerrando de ese modo el ciclo.**

en otro componente del sistema distinto de los átomos de manganeso. Uno de nosotros (Govindjee), junto con Subhash Padhye, Takeshi Kambara y David N. Hendrickson, de la Universidad de Illinois en Urbana-Champaign, propuso en 1986 que el aminoácido histidina de una de las proteínas del reloj podría almacenar tal vez una carga positiva.

Los trabajos de Melvin P. Klein, Kenneth Sauer y sus colaboradores, de la Universidad de California en Berkeley, y de Robert R. Sharp y sus colegas, de la de Michigan en Ann Arbor, han contribuido a definir con mayor precisión los estados de oxidación de algunos de los átomos de manganeso. De forma aproximada, $S_0$ se ha identificado con la presencia de Mn(II), $S_1$ con la de Mn(III) y $S_2$ con la de Mn(IV). Tanto Mn(II) como Mn(III) parecen ser estables y de larga duración en el fotosistema II; estas observaciones corroboran la predicción de Kok, en virtud de lo cual los estados $S_0$ y $S_1$ serían estables. Por contra, el estado Mn(IV) asociado con $S_2$ es un intermediario bastante efímero. Las pruebas recientes obtenidas en el laboratorio de Horst T. Witt, de la Universidad Técnica de Berlín Oeste, indican que un ion Mn(II) se transforma en Mn(III) durante la transición $S_0$-$S_1$. Las únicas transformaciones observadas durante las siguientes transiciones son las de Mn(III) a Mn(IV).

Los estudios de RPE a baja temperatura llevados a cabo por G. Charles Dismukes y Yona Siderer, de la Universidad de Princeton, sugieren que los estados $S_2$ y $S_3$ implican complejos multinucleares con al menos cuatro átomos de manganeso. Así, el estado $S_2$ podría ser un grupo de valencias mixtas compuesto por un átomo de Mn(III) y otro de Mn(IV) o por tres átomos de Mn(III) y uno de Mn(IV).

En resumen, los cambios dinámicos de los estados de oxidación de los átomos de manganeso unidos al fotosistema II se corresponden indudablemente con cambios en los estados S del reloj de Kok. Cuál sea la configuración química y electrónica precisa de estos estados sigue siendo objeto de estudio.

Varios experimentos han indicado que es probable que el manganeso no se una directamente a ninguno de los pequeños polipéptidos del complejo del fotosistema II, lo que sugiere que los grandes polipéptidos D1 y D2 son los sitios más problables de unión del manganeso. Recientemente hemos propuesto que pueden existir cuatro sitios de unión del manganeso con los polipéptidos D1 y D2 en la cara interna de la membrana tilacoidal, pero otros investigadores han sugerido que el manganeso se fija en la zona de unión entre los polipéptidos D1, D2 y el de 33 kilodalton.

Los estudios de espectroscopía de rayos X realizados por Klein, Sauer y sus colaboradores, en Berkeley, y por Graham N. George y Roger C. Prince, en el Centro de Investigación Exxon en Annandale, han revelado algunos detalles sobre la disposición de los átomos de manganeso. En el estado $S_1$, dos de los átomos parecen formar parte de un complejo binuclear; están separados entre sí sólo 2,7 angstrom. (Un angstrom es la diezmillonésima parte de un milímetro.) El otro par de átomos de manganeso está separado por una distancia mayor. Podríamos asociar los átomos a los cuatro vértices de un trapezoide.

Gracias a estos estudios, se conoce ahora mucho mejor el modo en que los átomos de manganeso podrían catalizar la eliminación de electrones del agua para reducir $P680^+$. Los electrones, sin embargo, no son los únicos protagonistas de nuestra historia: la reacción de rotura del agua produce también cuatro protones. ¿Se liberan los cuatro protones a la vez, simultáneamente con el desprendimiento de oxígeno, o se liberan de manera secuencial junto con los electrones?

Esta cuestión ha sido resuelta mediante detalladas mediciones de liberación de protones en respuesta a una serie de destellos luminosos. Debido a que los protones aumentan la acidez del líquido circundante, la cronología de liberación de protones se puede abordar con electrodos y sustancias colorantes muy sensibles a la acidez. C. Frederick Fowler, de Martin Marietta, y, poco después, Satham Saphon y Anthony R. Crofts, entonces en la Universidad de Bristol, descubrieron que los cuatro protones se liberaban en secuencia: uno durante la transición $S_0$-$S_1$, ninguno durante la $S_1$-$S_2$, uno durante la $S_2$-$S_3$ y dos durante la transición $S_3$-$S_4$-$S_0$.

Estos hallazgos tienen importantes implicaciones para el mecanismo del reloj oxidante del agua, aunque su interpretación depende de que los protones liberados provengan directamente de las moléculas de agua o de alguna otra fuente, como los polipéptidos que unen los átomos de manganeso. Si los protones proceden del agua, las moléculas de ésta deben sufrir ciertos cambios químicos antes de llegar al estado $S_4$. Por contra, si los protones liberados en secuencia proceden directamente de los polipéptidos (y se sustituyen luego por protones de las moléculas de agua), la oxidación del agua no debe ocurrir hasta la transición final $S_4$-$S_0$. Ignoramos todavía el origen inmediato de los protones.

Con independencia de la fuente de protones, parece probable que los estados S superiores (en particular $S_2$) acumulen alguna carga neta positiva. Es posible, por tanto, que sea necesario un ion cargado negativamente para estabilizar esta carga positiva, lo que explicaría la observación de que iones como el cloruro sean esenciales para mantener en funcionamiento el reloj oxidante del agua. Seikichi Izawa, de la Universidad estatal de Wayne en Detroit, fue uno de los primeros en demostrar que los iones cloruro pueden poner en marcha el reloj oxidante del agua.

En 1982, en colaboración con Herbert S. Gutowsky y sus colegas de la Universidad de Illinois en Urbana-Champaign, empezamos a aplicar las técnicas de RMN al análisis de la unión de los iones cloruro a las membranas fotosintéticas. En estudios anteriores, Ion C. Baianu, Christa Critchley y uno de nosotros (Govindjee) observamos que los iones cloruro se asociaban y disociaban libre y rápidamente de membranas aisladas de cloroplastos. Estos resultados nos lle-

varon a especular en 1983 sobre la posible relación entre la unión de un ion cloruro cargado negativamente y la llegada de una carga positiva desde $P680^+$ al reloj oxidante del agua, y la verosímil coincidencia entre la liberación del ion cloruro y la liberación de protones.

Los experimentos de RMN realizados por Christopher Preston y R. J. Pace, de la Universidad Nacional de Australia en Canberra, sugirieron que los iones cloruro se unían más fuertemente en los estados $S_2$ y $S_3$ que en los estados $S_0$ y $S_1$. Este resultado coincide, en principio, con la mayor carga positiva de los estados S superiores. Los datos de espectroscopía de rayos X obtenidos por Klein indican que el cloruro no se une directamente a los átomos de manganeso en los estados S iniciales.

Peter H. Homann y sus colegas, de la Universidad estatal de Florida, han sugerido que el cloruro probablemente se une a aminoácidos dotados de carga positiva de las proteínas del reloj. En colaboración con Gutowsky hemos estudiado la unión del cloruro en complejos del fotosistema II de espinaca. Nuestros resultados indican que varios iones cloruro se unen al reloj y parecen distribuirse entre dos sitios principales de unión: uno cerca del manganeso, quizás en los polipéptidos D1 y D2, y el otro en el polipéptido de 33 kilodalton.

Estos experimentos sugieren que la función de los iones cloruro en el reloj oxidante del agua podría ser la de facilitar la liberación de los protones del agua. Con ello, los iones cloruro aumentarían el rendimiento de las reacciones de oxidación del agua, estabilizarían los iones manganeso cargados en los estados S superiores o ambas cosas. El papel del cloruro, aunque todavía objeto de controversia, podría ser el de organizar las proteínas del fotosistema II en una estructura estable.

Otro ion, el calcio ($Ca^{2+}$), es esencial tanto para la oxidación del agua como para la operación del centro de reacción del fotosistema II, y su función parece hallarse también estrechamente ligada a la del cloruro. Los experimentos de varios laboratorios sugieren que los iones calcio pueden remplazar en su función a dos de los polipéptidos de la parte inferior del fotosistema II implicados en la producción de oxígeno molecular. Se ha observado, asimismo, que la eliminación de los iones de calcio parece bloquear no sólo el reciclaje del reloj oxidante del agua (impidiendo la transición $S_3$-$S_4$-$S_0$), sino también la reducción rápida de $P680^+$ a P680.

Parece probable, por tanto, que el calcio tenga una función estructural o reguladora del fotosistema II. Se ha demostrado que el calcio desempeña un importante papel en el control de una amplia variedad de proteínas en otros sistemas biológicos: activa y desactiva las proteínas y mantiene su estructura tridimensional. Quizás en el fotosistema II los iones de calcio sitúen a los polipéptidos del reloj oxidante del agua en la conformación funcional adecuada.

El complejo mecanismo productor de oxígeno durante la fotosíntesis representa sólo una pequeña parte del proceso completo que desarrollan los organismos fotosintéticos oxigénicos. Aunque los aspectos generales son similares en todas las especies fotosintéticas, en el curso de la evolución se han ido produciendo diferencias significativas.

La mayoría de los análisis indica que las diferencias entre el fotosistema II de las cianobacterias y el de las plantas son relativamente pequeñas, lo que sugiere que las cianobacterias son antepasados de las plantas o, en cualquier caso, están estrechamente emparentadas. Las diferencias entre el centro de reacción de las cianobacterias y el de muchas otras bacterias fotosintéticas son bastante más pronunciadas, lo que indica una clara divergencia en la ruta evolutiva. La investigación continuada de los fotosistemas mediante genética molecular, espectroscopía y cristalografía de rayos X perfeccionará obviamente nuestros conocimientos sobre la evolución de la vida.

BIBLIOGRAFIA COMPLEMENTARIA


RELEVANCE OF THE PHOTOSYNTHETIC REACTION CENTER FROM PURPLE BACTERIA TO THE STRUCTURE OF PHOTOSYSTEM II. Harmut Michel y Johann Deisenhofer en *Biochemistry*, vol. 27, n.º 1, págs. 1-7; 12 de enero de 1988.

MECHANISM OF PHOTOSYNTHETIC WATER OXIDATION. Gary W. Brudvig, Warren F. Beck y Julio C. de Paula en *Annual Review of Biophysics and Biophysical Chemistry*, vol. 18, Annual Reviews, Inc., 1989.

PHOTOSYSTEM II: FROM A FEMTOSECOND TO A MILLISECOND. Govindjee y Michael R. Wasielewski en *Photosynthesis*. Dirigido por Winslow R. Briggs. Alan R. Liss, Inc., 1989.

PHOTOSYSTEM II, THE WATER-SPLITTING ENZYME. A. W. Rutherford en *Trends in Biochemical Sciences*, vol. 14, págs. 227-232; junio de 1989.

# Avances en semiconductores de arseniuro de galio

*Este compuesto no es un candidato para sustituir al silicio. Sin embargo, su velocidad y propiedades ópticas han dado lugar a muchas aplicaciones en el campo de los ordenadores y las comunicaciones optoelectrónicas*

Marc H. Brodsky

La era de la electrónica se asienta sobre la explotación de circuitos microscópicos grabados sobre láminas de silicio ("chips"). El éxito actual y las perspectivas futuras del silicio en los sistemas electrónicos domésticos, comerciales, industriales y militares explica la actitud crítica de quienes trabajan con este material ante la aparición de otro semiconductor prometedor, el arseniuro de galio. "El arseniuro de galio", dicen con sorna, "es la tecnología del futuro, siempre lo ha sido y siempre lo será".

Al cabo de casi treinta años de ser la tecnología del futuro, el arseniuro de galio ha empezado a abrirse su propio camino, no para desbancar al silicio, sino para servirle de complemento en nuevas aplicaciones. Las ventajas específicas de este material se basan en la velocidad con que los electrones se mueven a través del mismo, en operaciones con señales de baja intensidad y en la generación y detección de la luz. Estas ventajas revisten especial interés en numerosas aplicaciones relacionadas con los ordenadores, la recepción de señales de televisión y la transmisión optoelectrónica de datos a través de redes de fibras ópticas (la fotónica, como se denomina esta técnica). Los láseres y diodos emisores de luz de arseniuro de galio, que se utilizan en los equipos de reproducción de discos de audio y en las técnicas de producción de imágenes visuales, representan actualmente una cifra de ventas de más de mil millones de dólares anuales. Cada año se venden cientos de miles de antenas receptoras de las señales procedentes de satélites en las que se utilizan detectores de arseniuro de galio. Se prevé que el empleo de circuitos de alta velocidad que utilicen transistores de arseniuro de galio alcanzará una difusión similar dentro de muy pocos años. En una economía y una sociedad que dependen del intercambio rápido de información y de su procesado subsiguiente, se exigirá que muchos de los dispositivos basados en el silicio incorporen una proporción considerable de componentes de arseniuro de galio para realizar adecuadamente su trabajo.

MARC H. BRODSKY ha dirigido muchos trabajos de investigación sobre las propiedades fundamentales y las aplicaciones de los semiconductores, en el Centro de Investigación Thomas J. Watson de la empresa IBM en Yorktown Heights. Estuvo al frente del laboratorio de tecnología avanzada de arseniuro de galio hasta su reciente nombramiento como director de planificación técnica en la división de investigación de esta empresa. Brodsky cursó todos sus estudios de física en la Universidad de Pennsylvania.

La tecnología del arseniuro de galio ha seguido las huellas del curso de desarrollo que los científicos trazaron en su día para el silicio. Desde la invención del transistor, en 1948, por John Bardeen, Walter H. Brattain y William B. Shockley, de la compañía Bell Telephone, los investigadores han intentado mejorar los semiconductores siguiendo dos caminos diferentes. En primer lugar, los físicos e ingenieros electrónicos buscan materiales que sean capaces de conectar y desconectar un circuito con la mayor rapidez posible, y que lleven a cabo además otras operaciones, tales como la detección y producción de luz. La verdad es que la preparación del arseniuro de galio —un material que no existe en la naturaleza— se llevó a cabo, en los años cincuenta, gracias a Heinrich Welker, de los Laboratorios Siemens, con la intención de cubrir estos objetivos. Welker investigó, con idéntico pormenor, otros semiconductores relacionados con este material, derivados de elementos situados en las columnas del sistema periódico adyacentes a la del silicio y germanio, que fueron los elementos constituyentes de los primeros transistores.

En segundo lugar, los ingenieros se han preocupado por mejorar las técnicas de fabricación de los semiconductores. Este trabajo exige que las propiedades físicas y químicas de los semiconductores estén bien definidas y que se desarrollen los procesos y los materiales auxiliares compatibles con ellos, para fabricar los aislantes, conductores, conexiones externas y otros componentes esenciales de los correspondientes circuitos. Los materiales semiconductores han de purificarse con sumo cuidado, combinarse con otras sustancias en proporciones precisas y producirse en forma de cristales perfectos; las imperfecciones originadas durante el proceso de fabricación de los transistores y circuitos deben eliminarse sin comprometer las propiedades electrónicas deseadas. No es fácil el cumplimiento, ni siquiera hoy, de ninguno de estos requisitos; mucho menos lo fueron en un principio, cuando tuvo que crearse una nueva ciencia de los materiales conjugando una serie de estudios fundamentales en física, química, metalurgia y otras disciplinas. Intentaré guiar al lector a través de estos campos atractivos, e interrelacionados, de la física, la ingeniería, los materiales y la electrónica, para demostrar por qué el arseniuro de galio es un material cuya comercialización constituye un reto y se nos ofrece, a la vez, tan prometedora.

La propiedad más interesante del

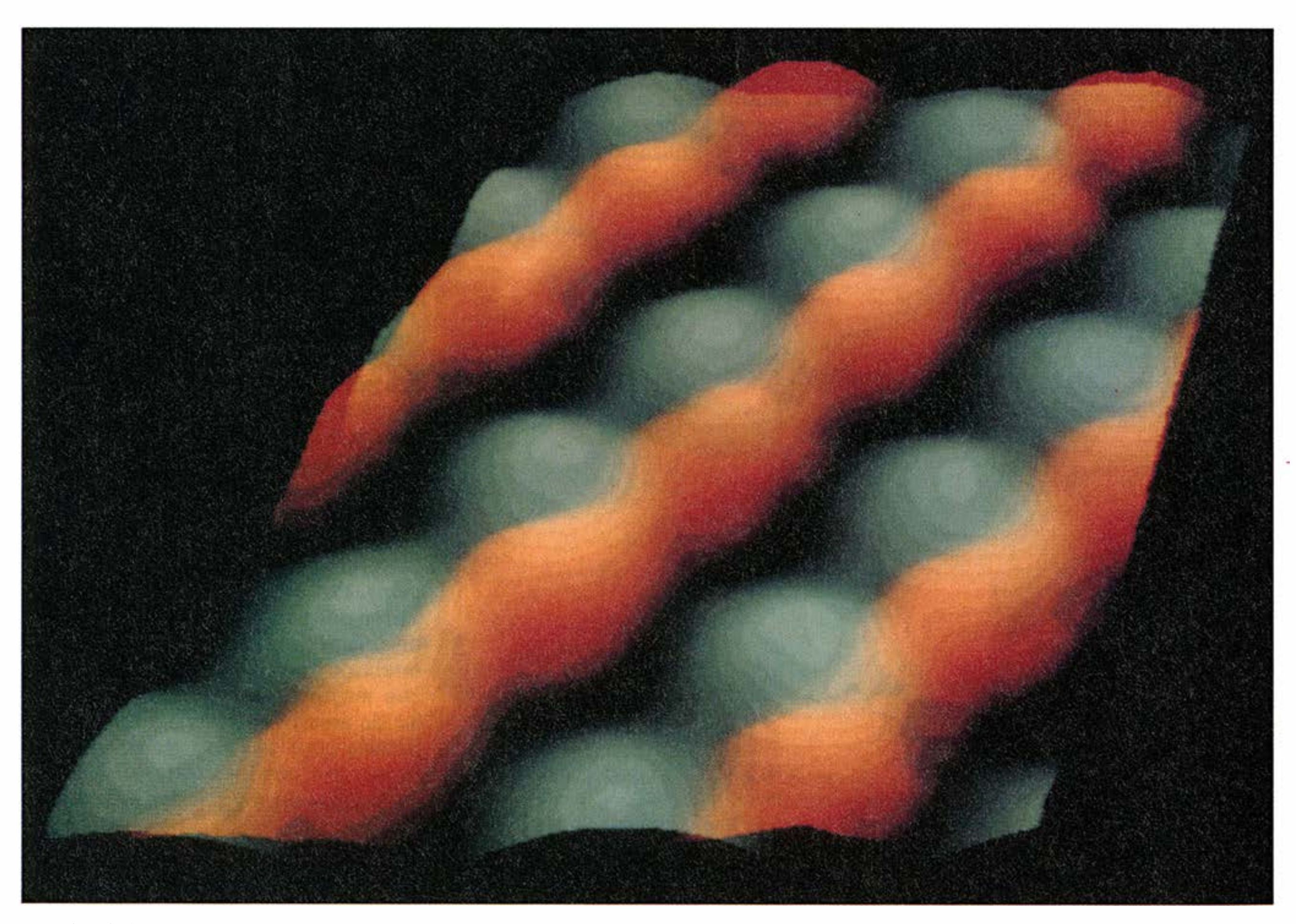
arseniuro de galio es la gran facilidad con que los electrones se mueven en su seno: en identidad de condiciones, los circuitos de arseniuro de galio son más rápidos que los circuitos de silicio para potencias iguales o incluso inferiores. Debido a que el arseniuro de galio consume menos energía, la producción de calor que ha de extraerse del circuito resulta menor. Esta propiedad adquiere destacada importancia, habida cuenta de la relación existente entre la velocidad de un semiconductor y la energía que consume.

Un ingeniero debe prestar atención al problema de la velocidad en el contexto de un dispositivo concreto, y no en un cristal puro de un determinado elemento o compuesto. En la actualidad, se utilizan diferentes tipos de transistores como elementos de conexión y desconexión en circuitería electrónica [*véanse las figuras 2 y 3*]. La realización de cálculos o el procesamiento de datos se llevan a cabo efectuando ciertos cambios en estos dispositivos. Estos cambios no pueden ser más rápidos que la velocidad de conexión o desconexión, es decir, el tiempo que un electrón invierte para atravesar la región del semiconductor que se encuentra bajo el control de las señales eléctricas procedentes de otra parte del circuito.

La velocidad de conexión o desconexión de un semiconductor depende de la velocidad media con que un electrón se mueve en su seno; esta velocidad es del orden de un millón o más de centímetros por segundo. En su movimiento a través de un transistor, los electrones encuentran numerosos obstáculos; como consecuencia de un gran número de colisiones con los átomos, iones u otros electrones presentes, los electrones en cuestión adquieren una distribución de velocidades característica, cuya forma viene determinada por el campo eléctrico bajo el que se mueven y por la manera en que resultan difundidos por los átomos que constituyen el semiconductor y las impurezas existentes en el mismo. Los electrones se mueven en todas direcciones y pierden, a menudo, una parte de su energía, disminuyendo su flujo neto en la dirección del campo eléctrico aplicado.

Una analogía mecánica nos ayudará a explicar de qué modo las propiedades físicas influyen sobre el movimiento de un electrón en un semiconductor. Podemos representar dos materiales semiconductores diferentes mediante sendos tubos, alineados con otros tantos obstáculos vibrantes y estacionarios e inclinados bajo un mismo ángulo respecto a la horizontal. Dichos obstáculos simbolizan el mecanismo de difusión, la inclinación proporciona un campo gravitatorio que corresponde al campo eléctrico y las esferas que caen por el interior de los tubos indican los electrones. En este símil, la velocidad de conexión equivale al tiempo que invierten las esferas en alcanzar el extremo inferior de los tubos. Por otra parte, la movilidad de los electrones viene representada por la facilidad con que las esferas salvan los obstáculos; esta cualidad viene reflejada por el tamaño de las esferas: cuanto menores sean, menos colisiones se producirán con los obstáculos.

En un semiconductor, los electrones se mueven a través de un conjunto de átomos constituyentes que se disponen en una red cristalina deter-



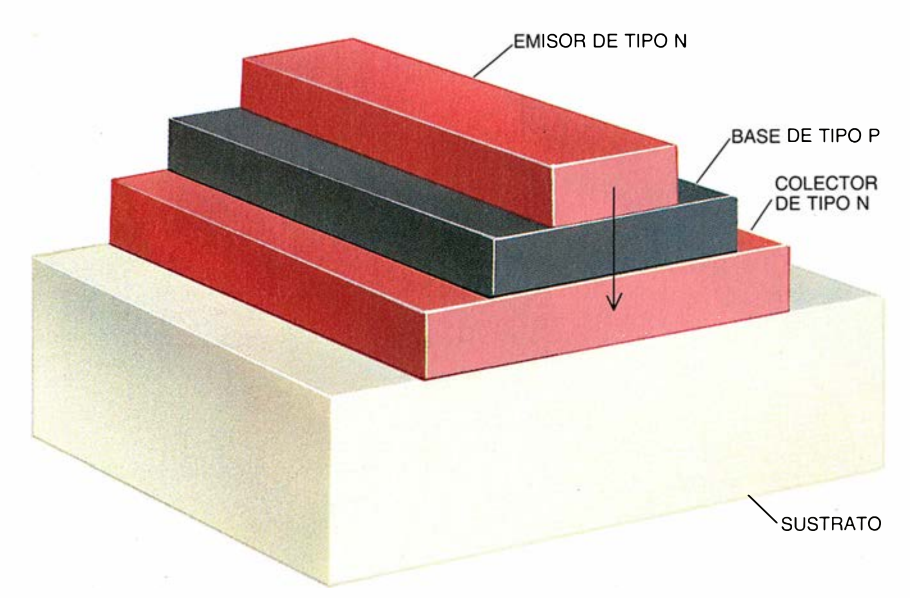
**1. DISPOSICION ANIDADA de los átomos de arsénico (*naranja*) y de galio (*verde*) que constituyen el arseniuro de galio, en una microfotografía obtenida por Randall M. Feenstra y Joseph A. Stroscio, de la compañía IBM. La imagen recoge los contornos electrónicos de los átomos citados.**

minada. Dado que los electrones de conducción se comparten entre todos los átomos presentes, dicha red tiene el carácter electrónico de un simple tubo respecto al paso de los electrones. Las estructuras formadas por átomos de galio y de arsénico que se pueden observar en la imagen de microscopía electrónica que reproducimos en la figura 1 atraen a los electrones en movimiento con menor intensidad de lo que lo hacen las estructuras de átomos de silicio. En consecuencia, los físicos consideran que la masa efectiva de los electrones resulta menor en el arseniuro de galio que en el silicio. Dado que otros factores contribuyen también a proporcionar una mayor movilidad, los electrones pueden alcanzar mayores velocidades en el arseniuro de galio que en el silicio; la distancia recorrida entre dos colisiones consecutivas resulta ser también mayor.

La difusión se puede evitar por completo acortando los recorridos críticos por debajo de la distancia media entre las colisiones sucesivas de los electrones. Esta circunstancia aseguraría que la mayoría de los electrones atraviesen la zona crítica realizando recorridos balísticos [véase "Electrones balísticos en semiconductores", por Mordehai Heiblum y Lester F. Eastman, INVESTIGACIÓN Y CIENCIA, abril de 1987]. Sin embargo, no se espera alcanzar la aplicación práctica de este principio hasta que transcurran algunos años desde el momento en que los transistores de arseniuro de galio convencionales encuentren su lugar en el menú tecnológico.

La analogía de las esferas y tubos con las colisiones de los electrones resulta aplicable sólo a las aceleraciones que se producen bajo la acción de campos eléctricos cuya intensidad esté por debajo de 10.000 volt por centímetro. La aplicación de campos superiores origina interacciones físicas adicionales que producen saturación, situación en la que los electrones se hacen realmente más pesados; en este caso, la ventaja del arseniuro de galio sobre el silicio se reduce o incluso se invierte. De hecho, el arseniuro de galio tiende a trabajar a voltajes óptimos más bajos que el silicio; esta circunstancia dificulta la conexión de circuitos basados en estos dos tipos de semiconductores. Para que coexistan los dos semiconductores, se impone adoptar un voltaje de compromiso.

La movilidad importa también en las operaciones a alta frecuencia y bajo nivel de ruido. El ruido (fluctuaciones al azar en el voltaje que determinan el valor mínimo de la señal a utilizar) se puede reducir, para frecuencias elevadas, aumentando al máximo la movilidad de los electrones, tanto en el propio transistor, como en sus conexiones al resto del circuito. La utilización de circuitos de arseniuro de galio de bajo nivel de ruido adquiere notable valor en la detección de señales de televisión y microondas.




**2. ESTE TRANSISTOR BIPOLAR sirve para conectar un circuito con su emisor y su colector, dos regiones en las que existen electrones suministrados por dadores (contaminación de tipo *n*). Estas zonas están separadas por una región controlada, la base, que contiene algunos agujeros (contaminación de tipo *p*). Cuando a esta base se aplica una corriente, los electrones inyectados por el emisor atraviesan el transistor y llegan hasta el colector, conectando el circuito. Estos dispositivos se fabrican sobre obleas de silicio, utilizando procedimientos fotolitográficos y otros métodos de crecimiento y procesado. El arseniuro de galio que se utiliza en dispositivos bipolares se deposita formando tres capas diferenciadas; una o más de estas capas puede contener cantidades variables de aluminio.**

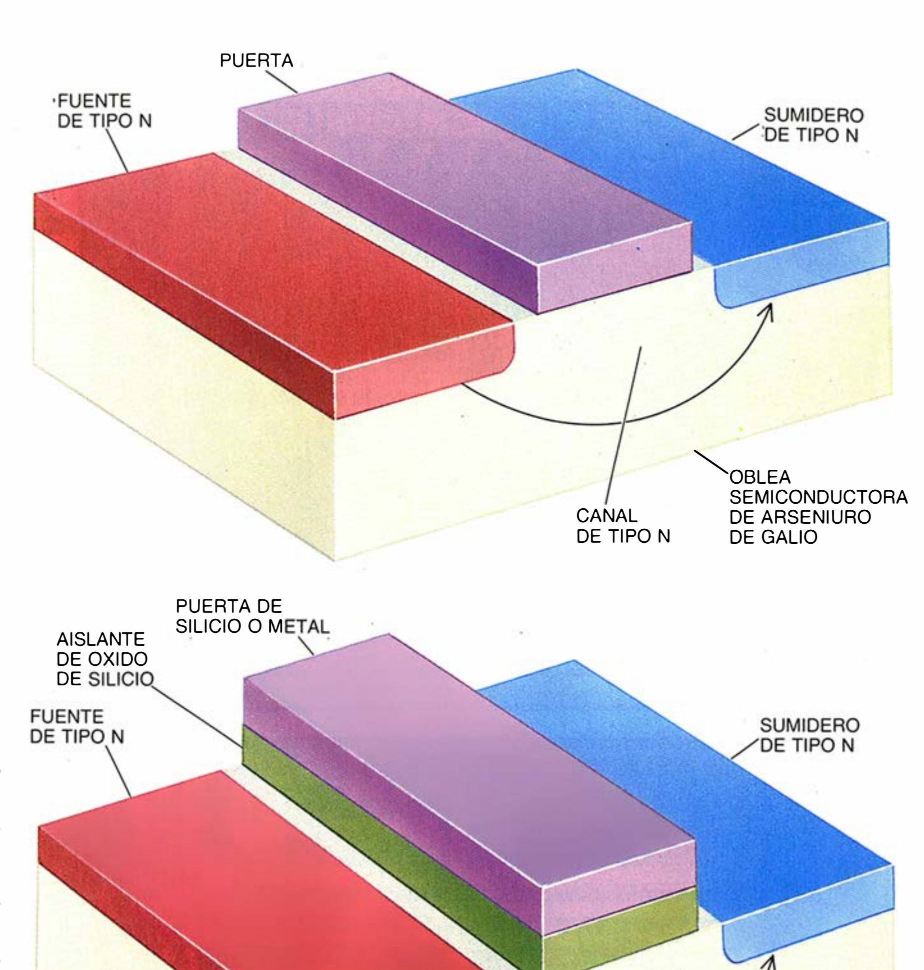
El otro motivo principal de la superioridad del arseniuro de galio sobre el silicio estriba en la facilidad, muy superior, con que pueden modificarse las separaciones entre sus bandas electrónicas o niveles de energía. Esta "ingeniería del intervalo entre bandas" puede originar ciertas propiedades optoelectrónicas muy versátiles y permite diseñar transistores más flexibles. Una banda electrónica define el intervalo de energías que posee un electrón y representa una forma ensanchada del estado energético de un electrón en un átomo. En un semiconductor puro, la banda de valencia (que contiene a los electrones responsables del enlace químico) está fundamentalmente llena o completa; el nivel de energía mayor inmediatamente siguiente, la banda de conducción, está fundamentalmente vacía. Mediante un proceso de contaminación, se crean algunas cargas móviles en las bandas; este proceso consiste en una adición cuidadosa de trazas de impurezas en el semiconductor de base. Las regiones que poseen electrones en la banda de conducción reciben el nombre de semiconductores de tipo *n*; se preparan contaminando el material base con átomos que se comportan como dadores de electrones; las regiones de tipo *p* se forman añadiendo agujeros cargados positivamente (cada uno de estos agujeros consiste, en realidad, en la ausencia de un electrón en la banda de valencia).

La diferencia de energías entre la parte superior de la banda de valencia y la parte inferior de la banda de conducción se denomina intervalo o separación entre las bandas. Este intervalo es mayor en el arseniuro de galio que en el silicio, pero su anchura puede aumentarse o reducirse sustituyendo los átomos que constituyen este material por otros diferentes (aleación). Por ejemplo, si el galio se sustituye por aluminio, para dar lugar al arseniuro de aluminio, la anchura del intervalo entre las bandas resulta mucho mayor. Las sustituciones parciales de átomos producen intervalos de bandas cuya anchura guarda directa proporción con la cuantía de aluminio en la aleación. Se pueden crear otras aleaciones valiosas sustituyendo algunos átomos de galio por indio, algunos átomos de arsénico por fósforo o recurriendo a ambas sustituciones a la vez.

Los espaciados interatómicos de las aleaciones de arseniuro de galio y aluminio se ajustan a los del arseniuro de galio puro hasta tal punto que ambos materiales pueden unirse átomo a átomo sin que aparezcan defectos. Se pueden utilizar capas muy delgadas de dos o más aleaciones alternadas para crear heterouniones, es decir, estructuras cuyos intervalos entre bandas varían de una capa a otra. Tenemos un ejemplo de esta posibilidad en la superred de semiconductor, inventada por Leo Esaki y sus colaboradores en la empresa IBM. Esta estructura singular consiste en un depósito de una serie de capas sucesivas y alternas de arseniuro de galio y aluminio y arseniuro de galio, depositadas sobre un sustrato de arseniuro de galio. Los electrones que se mueven paralelamente a los planos de las capas que constituyen esta especie de "bocadillo" resultarán confinados normalmente en el intervalo de bandas más bajo de las capas de arseniuro de galio. Para que estos electrones se muevan perpendicularmente a los planos de las capas o láminas, deben penetrar y atravesar las barreras de los intervalos de bandas del arseniuro de galio y aluminio. Variando el número, la anchura y la composición de dichas capas, se puede manipular las propiedades físicas y electrónicas de las heterouniones semiconductoras.

El crecimiento de cristales perfectos constituye un requisito esencial para aprovechar las ventajas que representa el intervalo de bandas variable del arseniuro de galio. Para nuestra desgracia, los espaciados atómicos de muchas aleaciones de arseniuro de galio que poseen las propiedades electrónicas adecuadas resultan inapropiados; sus redes cristalinas no se entremezclan entre sí, ni tampoco con las del arseniuro de galio. Si se procede a laminar dos cristales mal ajustados, algunas filas de átomos saltarán a través de las caras adyacentes, creando defectos que pueden propagarse a lo largo de la correspondiente capa y arruinar su aprovechamiento electrónico. Este efecto limita la lista de sustancias que pueden unirse. Se puede aceptar cierto desajuste entre las redes provocando que una de las capas cristalinas sea muy delgada, pero con ello restringimos también las posibilidades del diseño.

En algunos casos, se consiguen modificaciones importantes en la anchura del intervalo entre bandas utilizando un sustrato diferente. Por ejemplo, en los láseres de fosfuro y arseniuro de galio e indio se utilizan sustratos de fosfuro de indio; esta clase de láseres resulta muy indicada en las comunicaciones entre grandes distancias mediante fibras ópticas.



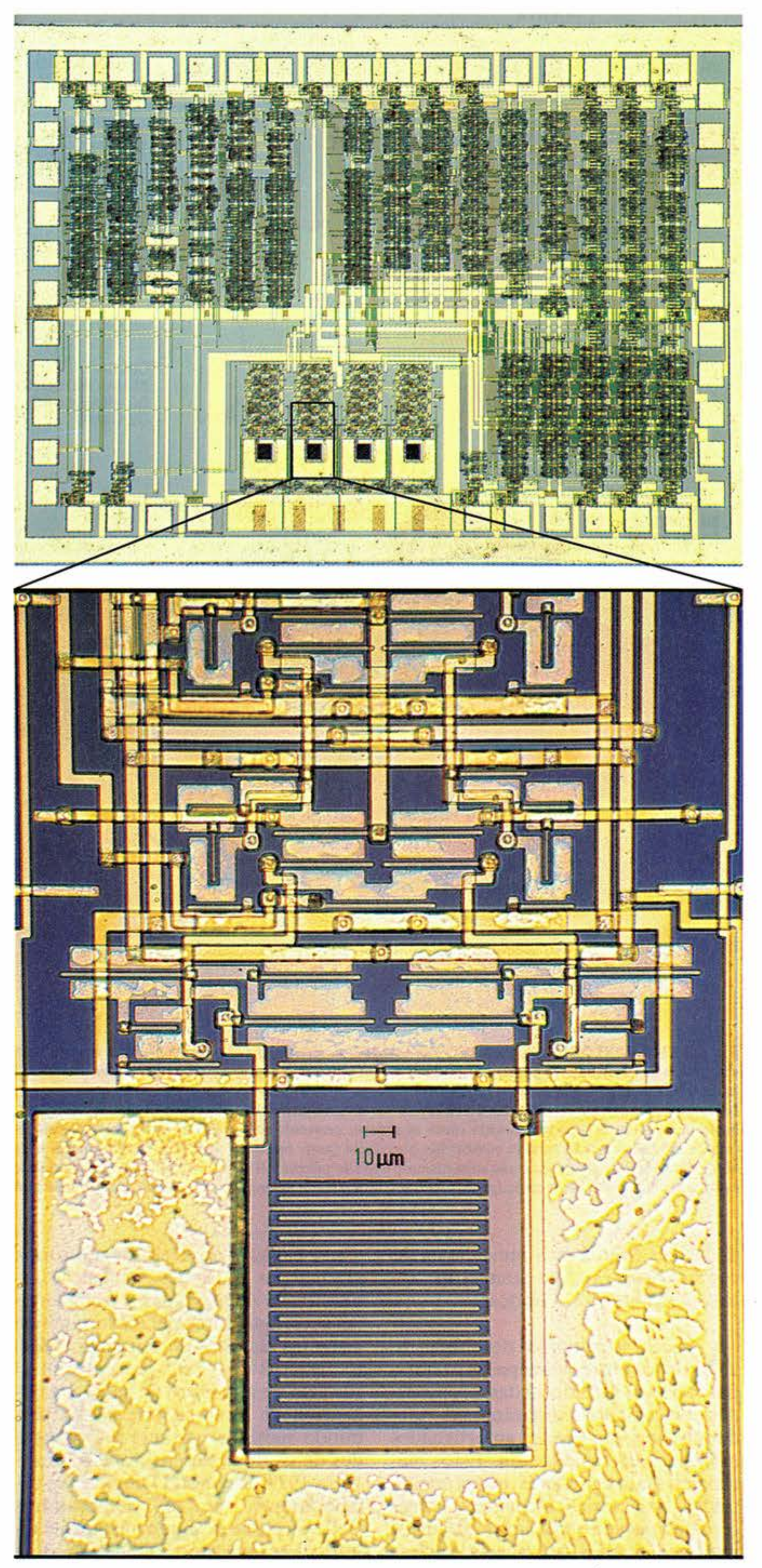

**3. LOS TRANSISTORES DE EFECTO DE CAMPO controlan el paso de electrones a través de un canal bajo un electrodo: la puerta. Cuando se aplica una carga positiva a ésta, atrae a los electrones hasta una región delgada situada en la parte superior del sustrato, creando un canal conductor con contaminación de tipo *n*. Este canal conduce a los electrones desde una región de tipo *n*, la fuente, hasta una región de tipo *n* conocida como sumidero, cerrando el circuito al que estos transistores están conectados. En la versión constituida por silicio (*parte inferior*), se utiliza el óxido natural de alta calidad de este elemento como capa aislante entre la puerta y el sustrato. Por contra, en el diseño ordinario con arseniuro de galio (*parte superior*), la puerta se encuentra ahora conectada al sustrato.**

A veces, se evitan los desajustes directos entre dos capas cristalinas que posean las adecuadas propiedades electrónicas separándolas mediante unas superredes amortiguadoras diseñadas con este fin. Estos materiales amortiguadores están constituidos por varias aleaciones cuyos cristales poseen espaciados atómicos intermedios a los de las capas activas y son capaces de absorber de forma gradual las tensiones mecánicas generadas en dichas capas. Los materiales amortiguadores permiten también aislar los circuitos de cualquier defecto residual que esté presente en el sustrato de arseniuro de galio. La utilización de estas superredes ha permitido mejorar bastante las cualidades de ciertos láseres y transistores.

En 1979, Raymond Dingle y sus colaboradores, de los Laboratorios Bell, pusieron a punto un nuevo diseño para transistores de heterounión

por efecto de campo, basado en la colocación de un canal de arseniuro de galio bajo una capa de arseniuro de galio y aluminio. Estos autores descubrieron que, cuando un átomo contaminante situado en una capa de arseniuro de galio cede un electrón a la banda de conducción, se convierte en un ion positivo que produce la dispersión de otros electrones. En consecuencia, procedieron a separar los átomos dadores de electrones presentes en el arseniuro de galio introduciendo ciertos contaminantes en una capa adyacente de arseniuro de galio y aluminio. Los electrones que proporcionan los átomos contaminantes se mueven hacia la banda de conducción más baja de la capa de arseniuro de galio más próxima y, por tanto, se mueven más deprisa de lo que lo harían si los átomos dadores ionizados permanecieran en el canal, bloqueando el camino de los electrones. Esta técnica de contaminación de modulación, así se llama, no tardó en incorporarse en los transistores de efecto de campo de arseniuro de galio fabricados por la japonesa Fujitsu y la francesa Thomson-CSF. Gracias a dicha técnica, se aumenta la movilidad de los electrones en un 20 por ciento a temperatura ambiente. Cuando se enfrían a temperaturas de 77 grados Kelvin (es decir, 77 grados centígrados por encima del cero absoluto), los transistores basados en la contaminación de modulación trabajan con una movilidad tres veces mayor que la de los dispositivos contaminados según las técnicas convencionales.

Otro aspecto material del arseniuro de galio que podría acelerar la velocidad de los circuitos es la capacidad que presentan las obleas fabricadas con este material para actuar como semiaislantes. Dichas obleas sirven de sustrato sobre el que se construyen los dispositivos y circuitos. Los sus-




**4. PASTILLA OPTOELECTRONICA de arseniuro de galio: recibe luz infrarroja sobre su fotodetector (*detalle*) a razón de mil millones de bits por segundo y la convierte en una señal eléctrica débil. El fotodetector y los circuitos de amplificación y procesado electrónico están montados en la misma pastilla. Un conductor cuya longitud resulta demasiado corta para recoger las señales espúreas procedentes de los circuitos próximos conecta el detector con el primero de una serie de amplificadores. Un conjunto de circuitos adicionales que contienen miles de transistores distribuyen los datos en paquetes de ocho bits (bytes) y los introducen a continuación en los ordenadores digitales. Esta pastilla forma parte de una unidad constituida por tres pastillas diferentes, capaces de recibir y transmitir datos con funciones optoelectrónicas sin precedente; este dispositivo ha sido fabricado en la unidad de investigación del arseniuro de galio de IBM.**

tratos de silicio presentan siempre cierta conductividad eléctrica, debido a que su intervalo entre bandas es menor, y debido a que las cantidades de contaminantes residuales activos (no deseados) es mayor; en consecuencia, su utilización introduce cierta capacitancia en los correspondientes circuitos, reduciendo la velocidad con que los electrones lo atraviesan. La mayor anchura del intervalo de bandas del arseniuro de galio permite preparar este material en forma semiaislante, bien manteniendo los sustratos completamente libres de contaminantes activos, bien incorporando ciertos contaminantes especiales autocompensantes para cancelar, casi del todo, el efecto de los contaminantes residuales. Estas ventajas sobre el silicio pueden desaparecer cuando los niveles de integración superan algunos miles de circuitos por pastilla, en cuyo caso la capacitancia entre los numerosos conductores de conexión de dichos circuitos, muy poco separados, determina la existencia de un límite crítico para la velocidad de la señal.

Además de la movilidad de los electrones y la flexibilidad del intervalo entre bandas, la tercera y más importante ventaja del arseniuro de galio sobre el silicio es su capacidad para emitir y detectar radiación en el infrarrojo próximo. En el arseniuro de galio, la energía potencial de un electrón que se mueva desde la banda de conducción hasta la de valencia puede emitirse en forma de un cuanto de radiación electromagnética o fotón. En el caso del silicio, esta misma reacción suele desarrollarse mediante un proceso no radiactivo, una colisión por ejemplo, con el fin de conservar el momento. Esta diferencia en las propiedades del intervalo entre bandas explica por qué el arseniuro de galio sustenta funciones optoelectrónicas y el silicio no.

Una fuente de radiación de arseniuro de galio consiste en un diodo en el que existen dos regiones contaminadas de forma opuesta; en la región de tipo *p*, existen muchos agujeros (en la banda de valencia) y en la región de tipo *n* existen muchos electrones (en la banda de conducción). La aplicación de un voltaje a este diodo *p*/*n* produce una inyección de agujeros y electrones en la zona que separa estas dos regiones, y crea allí un exceso de poblaciones de agujeros y electrones. Cuando un electrón y un agujero se combinan, ambos se aniquilan y la energía correspondiente al intervalo de bandas de los electrones se emite en forma de un fotón.

Esta recombinación electrón-agujero es extraordinariamente eficiente (bajo ciertas condiciones, es capaz de convertir la mayor parte de la energía eléctrica en energía luminosa) y puede utilizarse con gran facilidad para crear un componente electrónico familiar: el diodo emisor de luz (conocido por las iniciales de estas palabras en inglés, LED). Los LED de fósforo de arseniuro de galio o de arseniuro de galio y aluminio se encuentran en muchos aparatos electrónicos, bajo la forma de pequeñas luces de color amarillo o rojo, respectivamente. Cada año se fabrican por decenas de millones.

Los diodos de láser producen radiación más concentrada. Las caras del cristal que forman estos dispositivos son perfectamente paralelas entre sí y se comportan como espejos semitransparentes. La luz que se refleja en estas caras atraviesa una y otra vez la región de recombinación, donde estimula la emisión de radiación que tiene la misma longitud de onda y direccionalidad. El haz resultante es altamente coherente. Los láseres de diodo semiconductor se fabricaron por vez primera a partir de uniones *p*/*n* de arseniuro de galio en sendos ensayos llevados a cabo, de manera independiente en la IBM y General Electric, en 1962. Los diseños modernos, basados en el intervalo entre bandas, dependen de estructuras que contienen capas de composiciones diferentes. Por ejemplo, los láseres de arseniuro de galio emiten radiación cuyas longitudes de onda se sitúan en el infrarrojo próximo, debido a que el intervalo entre bandas del arseniuro de galio corresponde a una energía situada por debajo mismo de los fotones visibles. La mezcla del arseniuro de galio con otros elementos (aluminio, indio, fósforo) hace variar la anchura de dicho intervalo hasta valores mayores o menores, lo cual produce una salida que, según los casos, se sitúa en el infrarrojo más lejano o en la región visible.

La capacidad del arseniuro de galio y de sus aleaciones para detectar la luz, invirtiendo la reacción en la que se basa el funcionamiento de los LED y los láseres de diodos, reviste también interés. Los fotodetectores resultantes se pueden ajustar a una determinada longitud de onda por medio de las mismas técnicas basadas en el intervalo entre bandas que se utilizan para ajustar los láseres de diodos. Merced a su alto rendimiento, la respuesta de los fotodetectores de arseniuro de galio es mucho más rápida que la de los de silicio. Por otra parte, estos dispositivos presentan una ventaja adicional: su fácil integración en los circuitos de alta velocidad de arseniuro de galio.

La posibilidad de integración constituye un factor fundamental por razones económicas y funcionales. En primer lugar, como en todos los procesos basados en la miniaturización, la integración abarata los costes unitarios de estos dispositivos, pues se fabrican, a partir de una simple oblea, un gran número de pastillas ("chips") y cada pastilla contiene, a su vez, miles de circuitos. En segundo lugar, cuando se coloca un fotodetector muy cerca del circuito que produce la primera amplificación (el circuito del "extremo frontal"), se puede diseñar la conexión entre los dos elementos de forma tal que se reduzcan al mínimo las señales no deseadas que proceden de los circuitos próximos; estas señales constituyen un grave problema en los circuitos conectados a detectores no integrados mediante enlaces de conductores convencionales. Más allá de la integración de funciones electrónicas y ópticas, nos encontramos con la posibilidad más especulativa de fabricar pastillas en las que una señal óptica module directamente a otra. Aunque ya se han propuesto muchos dispositivos de este tipo, todavía estamos lejos de obtenerlos.

Además de la capacidad de producir y detectar luz, el arseniuro de galio posee otras muchas cualidades que resultan potencialmente interesantes para determinadas aplicaciones. Por ejemplo, su amplio intervalo de temperaturas a las que puede trabajar y su elevada resistencia a la radiación de alta energía le convierten en idóneo para la automoción y la industria militar, respectivamente.

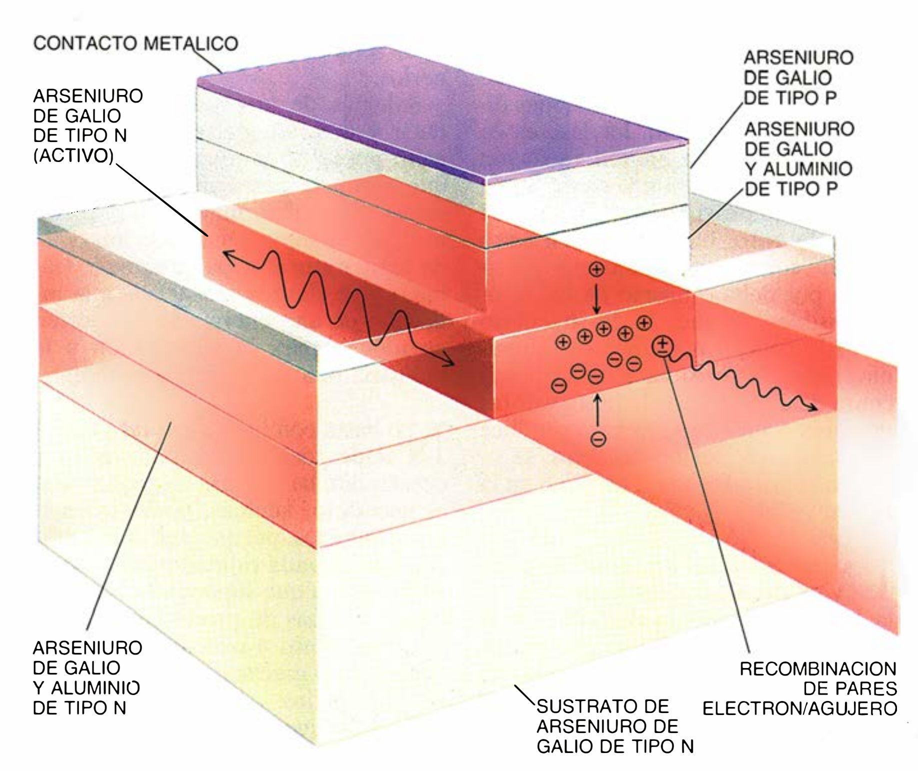
No basta con identificar las propiedades del arseniuro de galio y desarrollar las formas de utilizarlas; es necesario, además, poder fabricar los productos pertinentes y hacerlo con una elevada calidad y bajo coste. El problema que supone el desarrollo de las técnicas de producción, de las que nos vamos a ocupar a continuación, es, en muchos sentidos, el más difícil de todos.

Al igual que el silicio, el arseniuro de galio forma elementos electrónicos de acuerdo con ciertas condiciones impuestas por sus peculiares características físicas y químicas. Estas condiciones resultan agravadas, además, por el hecho de que muchos de los componentes de los circuitos integra-

dos no están constituidos por semiconductores. Entre estos componentes cabe citar a los conductores metálicos que conectan a los diferentes elementos, los aislantes que los separan, las resistencias y condensadores que controlan el flujo de la corriente y los contaminantes que proporcionan los electrones y los agujeros.

Uno de los principales inconvenientes del arseniuro de galio y de las aleaciones con él relacionadas es la inexistencia de un óxido natural utilizable, tal como el que el silicio forma cuando se calienta en el aire. El óxido de silicio forma un recubrimiento electrónico y mecánico que posee muchas aplicaciones. En los transistores de efecto de campo [*véase la figura 3*], el óxido de silicio proporciona el aislamiento necesario entre la puerta y el canal. En los circuitos de transistores bipolares, el óxido de silicio aporta el aislamiento entre los transistores adyacentes. El óxido de silicio se emplea también para la fabricación de transistores y microcircuitos; crea una capa protectora a través de la cual se pueden practicar algunas ventanas, para permitir que los reactivos químicos actúen y desarrollen ciertas figuras, depositar metales (para producir láminas conductoras) o añadir algunos contaminantes (y así activar determinadas regiones del material).

Por contra, en el caso del arseniuro de galio, los técnicos deben producir los aislantes y las estructuras de apantallamiento mediante otros medios, que, con frecuencia, resultan menos adecuados. Un diseño frecuente en los transistores de efecto de campo consiste en colocar una puerta metálica en contacto directo con el arseniuro de galio, lo cual origina un canal controlable mediante un efecto conocido como formación de una barrera de Schottky [*véase la figura 3*]. Otros diseños se basan en la utilización de capas cristalinas delgadas de arseniuro de galio y aluminio como barreras aislantes. Ahora bien, en tanto que el arseniuro de galio y aluminio presenta una gran flexibilidad para la contaminación modulada, tal como hemos explicado anteriormente, ni esta técnica ni la de las barreras de Schottky permiten alcanzar los valores de voltajes operativos que proporcionan los aislantes de óxido de silicio.




**5. LOS DIODOS DE LASER inyectan los agujeros procedentes de una capa de tipo *p* de arseniuro de galio y aluminio en una capa activa de arseniuro de galio. Una capa de tipo *n* de arseniuro de galio y aluminio inyecta los electrones desde abajo. El exceso de poblaciones de agujeros y electrones se aniquilan entre sí, y la diferencia de energías entre la banda de conducción de los electrones y la banda de valencia de los agujeros se libera en forma de un cuanto de luz. Los cuantos que se mueven en una dirección paralela a las caras verticales se reflejan hacia la capa activa, y su llegada a ésta estimula la emisión de cuantos que poseen la misma longitud de onda y direccionalidad que ellos.**

Otra desventaja muy importante del arseniuro de galio y de los compuestos relacionados con él se refiere al carácter compuesto de todos ellos. Mientras los defectos del silicio elemental se pueden eliminar con sólo calentar este material, a fin de agitar a sus átomos y obligarlos a que se alineen, en el arseniuro de galio este proceso entra en competencia con la vaporización selectiva del arsénico. Los defectos se producen en muchas de las fases de fabricación de los circuitos integrados. En particular, los contaminantes suelen incorporarse acelerando los iones correspondientes, con el fin de implantarlos en el semiconductor. En el caso del silicio, el daño que produce este proceso se puede reparar, y todos los contaminantes incorporados situar en los lugares adecuados del cristal, calentándolo hasta temperaturas de unos 1000 grados Celsius durante algunos minutos. (El recubrimiento de óxido de silicio impide que los contaminantes escapen del material durante esta operación.) En cambio, el arseniuro de galio con iones implantados no admite tal recocido, ni siquiera aunque la temperatura no supere los 800 grados Celsius. Las precauciones especiales adoptadas con el objeto de retener el arsénico y los contaminantes han alcanzado sólo un éxito parcial: entre un 90 y 95 por ciento de los contaminantes implantados originalmente resultan activados.

Se están ensayando nuevas técnicas basadas en ciclos de templado con una duración de segundos, y en capas de recubrimiento especiales (formando una estructura hermética); mas, por ahora, no existe una solución viable para llevar a cabo un templado completo del arseniuro de galio. Por cuya razón, las propiedades de estos dispositivos tienden a variar a lo largo de cada pastilla, y ello, junto con otros factores, impone ciertos límites prácticos sobre el nivel de integración que es posible alcanzar. En tanto que un microcircuito de silicio de un centímetro cuadrado puede mantener en la actualidad más de un millón de transistores, las pastillas de arseniuro de galio sólo pueden acomodar, en el mejor de los casos, algunas decenas de miles de componentes. Debido a que el coste de fabricación de una oblea de arseniuro de galio es el mismo o incluso superior que el de una oblea de silicio (cada una de las cuales puede contener muchos conjuntos de pastillas), este nivel de integración bajo constituye una seria desventaja. El tamaño relativamente menor de las obleas de arseniuro de galio que se

pueden fabricar en la actualidad ofrece una dificultad adicional desde el punto de vista económico: en general, dichas obleas suelen tener un diámetro comprendido entre siete y diez centímetros. En cambio, las obleas de silicio que se emplean tienen un diámetro comprendido entre doce y veinte centímetros.

Las aplicaciones actuales del arseniuro de galio se concentran en los dispositivos para los que el elevado coste unitario resulta soportable debido a que su función es única. Este semiconductor se utiliza, con mucha frecuencia, en la fabricación de receptores de alta velocidad, en los que se necesita una respuesta rápida y un bajo nivel de ruido, y en el campo de la generación óptica, para el que no existe un material sustitutivo. Las aplicaciones en la fabricación de los circuitos digitales instalados en los ordenadores modernos están abriéndose paso. Existe también el proyecto de usar arseniuro de galio en los futuros microprocesadores.

Quizá la aplicación más conocida de los detectores de arseniuro de galio se encuentre en la fabricación de los focos de antenas parabólicas. Los satélites de comunicación utilizan microondas con frecuencias de hasta 12 gigahertz, una región espectral en la que la velocidad del arseniuro de galio no tiene rival. El arseniuro de galio no sólo es capaz de convertir estas longitudes de onda en señales electrónicas claras, sino que puede también amplificar las débiles señales eléctricas iniciales sin introducir apenas ruido. Los amplificadores de silicio sólo pueden llevar a cabo esta tarea mediante conexiones de transistores que introducen un ruido bastante elevado que ahoga las débiles señales procesadas. Los avances experimentados en el proceso de integración deberían permitir la utilización futura de estos dispositivos en sintonizadores de televisión y otros productos comerciales; la viabilidad económica de esta posibilidad acaba de demostrarse en diferentes empresas de Japón y Europa. Se espera que estos sintonizadores proporcionen imágenes más claras en zonas donde la recepción de las señales plantea dificultades.

La aplicación del arseniuro de galio que hoy está recibiendo mayor atención pertenece al dominio de la transmisión fotónica de información. La propagación de la luz a través de fibras ópticas puede transportar mucha más información y llevarla mucho más lejos que lo que permiten las señales eléctricas en los conductores metálicos ordinarios. La interconexión de sistemas digitales y de otro tipo mediante fibras ópticas está avanzando con gran rapidez. Se están utilizando ya estas fibras en comunicaciones transcontinentales y transoceánicas, y su uso está llegando también a los sistemas telefónicos y de televisión por cable de nuestros hogares y oficinas. Y a las fibras ópticas se recurre para extender la distancia y velocidad de las conexiones entre ordenadores. Estas conexiones se convertirán paulatinamente en una parte del propio sistema de procesamiento digital. Algunos analistas creen que la importancia de la optoelectrónica para la sociedad del futuro será similar a la que poseen los actuales sistemas de cálculo digital.

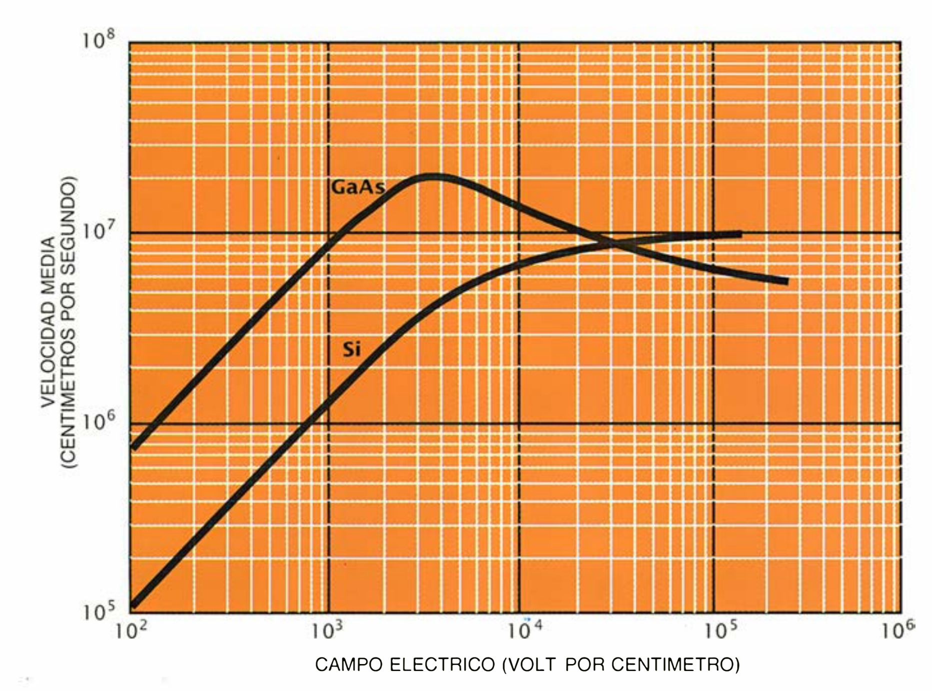
Los láseres de diodo representan frente a los láseres de gases lo que son los transistores respecto a las válvulas de vacío; la misma analogía se puede hacer extensiva cuando comparamos los LED con las bombillas incandescentes. En cada uno de estos casos, los semiconductores son mucho más pequeños, más eficientes y más económicos de fabricar y operar que sus contrapartidas. Las ventajas que presentan no se miden en términos de algunos puntos de porcentaje, sino en órdenes de magnitud. Estas diferencias hacen posible aplicaciones tan nuevas como el reproductor de discos compactos. Este producto tan familiar para los consumidores utiliza un láser de arseniuro de galio y aluminio en la lectura de los datos codificados en forma de marcas grabadas sobre un disco en rotación. Un láser de arseniuro de galio y aluminio que opera a una potencia mayor es un componente esencial de un dispositivo similar, una unidad de almacenamiento de datos de ordenador basada en un disco óptico.

Los láseres de estado sólido son candidatos particularmente apropiados para la ingeniería basada en el intervalo entre bandas. Esta técnica garantiza el control preciso de la longitud de onda que emite un láser. Los láseres obtenidos a partir de capas de fosfuro de indio, galio y aluminio depositadas en sustratos de arseniuro de galio emiten luz visible con una longitud de onda situada en la región del rojo. Se está ensayando la sustitución de estos dispositivos por láseres gaseosos, mucho más complicados, en lectores de los códigos de barras, que se usan para controlar los productos en las cajas de algunos comercios y los componentes de los automóviles en las líneas de montaje de las fábricas. Los láseres de fosfuro y arseniuro de indio y galio hallan generalizada aplicación en las comunicaciones a grandes distancias, porque puede ajustarse su salida a las longitudes de onda infrarrojas de 1,3 o 1,55 micrometros, que son las que se absorben menos fácilmente en las fibras ópticas.

Hasta ahora, la mayoría de las investigaciones en optoelectrónica se han concentrado en mejorar el funcionamiento de dispositivos específicos, y no en integrarlos en una unidad susceptible de ser incorporada en una pastilla. En parte, este planteamiento es la herencia del papel hegemónico desempeñado por las compañías de teléfono de largas distancias, que necesitan disponer de los enlaces ópticos adecuados en aplicaciones en las que el coste unitario es un factor de importancia secundaria, dado que su repercusión se reparte entre muchas conversaciones telefónicas. En consecuencia, láseres, detectores y transistores de los dispositivos comerciales utilizan pastillas separadas.

La conexión progresiva de ordenadores mediante fibras ópticas exigirá grandes cantidades de dispositivos optoelectrónicos mucho menos caros. Estos dispositivos se pueden fabricar a unos precios unitarios razonables, sólo cuando se utilizan técnicas de integración a gran escala. Los enlaces optoelectrónicos acabarán por transmitir datos a razón de mil millones de bits por segundo o incluso valores superiores; estas velocidades de transmisión se pueden conseguir, en principio, mediante transistores de arseniuro de galio bipolares o de efecto de campo. Los modernos transistores de silicio de diseño bipolar podrían también cubrir este objetivo. Sin embargo, los dispositivos de arseniuro de galio de efecto de campo han sido los elegidos, debido a que disipan menos energía y pueden extenderse más adelante para alcanzar velocidades de operación todavía más elevadas.

En 1979, Amnon Yariv y sus colaboradores del Instituto de Tecnología de California fabricaron el primer transistor y láser interconectados utilizando un microcircuito de arseniuro de galio. La empresa Matsushita y los Laboratorios Centrales de Investigación NEC, en Japón, han desarrollado también unos dispositivos similares basados en el fosfuro de indio; en muchos otros laboratorios se están realizando nuevas investigaciones con diferentes compuestos. El reto mayor consiste en superar los problemas que




**6. VELOCIDADES MEDIAS DE LOS ELECTRONES del arseniuro de galio. Quintuplican de lejos los del silicio bajo la acción de campos eléctricos cuya intensidad sea inferior a 1000 volt por centímetro.**

plantea la fabricación de las versiones óptimas de transistores y láseres en una sola pastilla.

Mis colegas del laboratorio de tecnología avanzada del arseniuro de galio y de otros laboratorios de la IBM, trabajando en tres centros diferentes del estado de Nueva York y en un centro de Suiza, han construido un trío de microcircuitos con capacidad para transmitir mil millones de bits por segundo. De esta velocidad no se perderá nada: la velocidad de las conexiones de transmisión decuplicará la celeridad de los ordenadores que conectan. Ello se conseguirá porque las conexiones transmiten los datos en serie, en tanto que los ordenadores procesan los datos en bloques de ocho bits, o bytes (contando los dos bits añadidos para comprobar la presencia de errores en la transmisión). Los ordenadores más rápidos producen ya un flujo de datos tal que excede, en mucho, la capacidad de transporte de los conductores de cobre para distancias superiores a los 200 metros (donde las señales eléctricas poco espaciadas entre sí comienzan a mezclarse). Los sistemas optoelectrónicos, con su anchura de banda considerablemente mayor, están remplazando a los cables para distancias cada vez más cortas; andando el tiempo, estos mismos sistemas transportarán información en el interior de los propios ordenadores.

Nuestro circuito consiste en un bloque de tres pastillas y un conjunto de fibras ópticas de conexión. La primera pastilla, un transmisor de arseniuro de galio, serializa los bytes que llegan a él a razón de 100 millones por segundo y los envía hacia la segunda pastilla, constituida por un conjunto de láseres de arseniuro de galio y aluminio.

Uno de sus cuatro láseres envía la señal a una fibra óptica, con una velocidad de mil millones de bits por segundo. Esta señal llega a la tercera pastilla, un receptor de arseniuro de galio, uno de cuyos cuatro fotodetectores la convierte en un impulso electrónico. Los impulsos alimentan directamente una serie de transistores de arseniuro de galio, que la amplifican [*véanse los detalles en la figura 4*]. Otros circuitos de la pastilla receptora deserializan la señal en bytes. En cada uno de los dos extremos del puente optoelectrónico, las pastillas del transmisor y del receptor, constituidas por arseniuro de galio, conectan este dispositivo con el circuito de silicio, que es el que procesa realmente los bytes como parte de un ordenador. Sin embargo, a medida que las velocidades del procesamiento de datos se incrementen, cabe esperar que el uso del arseniuro de galio se vaya extendiendo también a los circuitos digitales de los diferentes sistemas que están interconectados entre sí.

Esa circuitería constituye el corazón de los ordenadores digitales, un campo de extraordinaria importancia en el que el arseniuro de galio ha encontrado pocas aplicaciones hasta la fecha. Las aplicaciones digitales más comunes exigen la utilización de circuitos con niveles de integración más elevados (y, por tanto, con costes por unidad de circuito más bajos) que lo que permite el arseniuro de galio. La ventaja que representa la velocidad del arseniuro de galio sólo se ha explotado en ciertas aplicaciones digitales especiales de redes de ordenadores y superordenadores, es decir, en sistemas en los que interesan más las prestaciones que el coste. (Los superordenadores alcanzan velocidades de reloj de hasta 200 megahertz: una velocidad seis veces mayor que las de los ordenadores ordinarios más rápidos). Por otra parte, estas aplicaciones especializadas no resultan tan sensibles a los requisitos de miniaturización del silicio, dado que los transistores de silicio más potentes generan mucho calor y su densidad debe limitarse con gran cuidado. En consecuencia, los circuitos deben disponerse sobre una serie de pastillas diferentes, y esta exigencia crea un nuevo cuello de botella en los retrasos que se producen en la transmisión entre dichas pastillas. Los expertos en ingeniería informática intentan reducir al mínimo estos retrasos montando todas las pastillas relacionadas lo más cerca posible unas de otras.

El silicio continuará siendo el componente primordial de los ordenadores digitales hasta que con el arseniuro de galio se logren mayores niveles de integración y compatibilidad de empaquetamiento. Las demoras entre diferentes pastillas deberán también reducirse al mínimo, de forma tal que las ventajas que se alcancen en el dominio de una pastilla no se pierdan al pasar al dominio del sistema entero. Los proyectos de superordenadores existentes en Japón y Estados Unidos están utilizando unos niveles modestos de integración en los circuitos de arseniuro de galio y unas técnicas de empaquetamiento muy refinadas. En el superordenador Cray-3, el primero de su clase basado en el arseniuro de galio, se espera poder alcanzar velocidades de reloj de hasta 300 megahertz.

No todas las desventajas competitivas del arseniuro de galio son de tipo técnico. Aunque la tecnología del arseniuro de galio ha utilizado muchos de los métodos que se desarrollaron, en un principio, para el silicio, tales como la fotolitografía, esta ventaja ha sido compensada por las enormes inversiones que se realizaron con el silicio en el pasado. Esta situación explica la lentitud de los avances que ha experimentado el arseniuro de

galio, aunque su funcionamiento no sea mucho peor que el de aquél. Por otra parte, debido a que la mayoría de las nuevas inversiones que se empeñan en la investigación de semiconductores todavía se siguen canalizando hacia el silicio, nuestro material mantiene hoy un enorme atractivo. Las ventajas del arseniuro de galio en velocidad, bajo nivel de ruido y optoelectrónica deben ser valoradas en sus justos términos para que se pueda afrontar el reto de producirlo con un coste bajo y vencer así al silicio.

El camino tecnológico por el que el arseniuro de galio puede llegar a invadir, algún día, el campo del cálculo digital queda sugerido por un nuevo sistema de cálculo con instrucciones reducidas (conocido por las siglas RISC), una nueva técnica que incrementa las velocidades de cálculo en las pastillas utilizando un número reducido de elementos para llevar a cabo ciertas funciones especializadas. Diversas compañías han proyectado sistemas RISC de forma tal que las velocidades sólo pueden ser soportadas por el arseniuro de galio. Las pastillas de los sistemas RISC que existen están basadas en transistores de silicio que procesan unos 35 millones de instrucciones por segundo. La tecnología del silicio puede ya aceptar hasta tres veces este valor. Las velocidades superiores a 100 millones de instrucciones por segundo sólo se alcanzan con transistores de arseniuro de galio. Por razones de tipo económico, es necesario montar al menos 40.000 transistores en cada pastilla, un nivel de integración que ya ha sido alcanzado por diferentes compañías en la californiana "Quebrada del Galio"; se espera que, en el plazo de uno o dos años, estas compañías lancen al mercado los correspondientes diseños de circuitos plenamente comprobados. Quizás en ese momento, cuando los ordenadores, los enlaces entre ordenadores, los televisores y los discos compactos contengan todos ellos arseniuro de galio, podremos decir, al fin, que la tecnología del futuro se ha convertido ya en una realidad.


BIBLIOGRAFIA COMPLEMENTARIA

MOLECULAR BEAM EPITAXY AND HETEROSTRUCTURES. Dirigido por Leroy L. Chang y Klaus Ploog. Kluwer, 1985.

SEMICONDUCTOR DEVICES, PHYSICS AND TECHNOLOGY. S. M. Sze. John Wiley & Sons, Inc., 1985.

IEEE GaAs IC [GALLIUM ARSENIDE INTEGRATED CIRCUITS] SYMPOSIUM. IEEE, 6-9 de noviembre de 1988.

# Así comparten su alimento los vampiros

*Dos noches sin su ración de sangre y el vampiro muere de inanición, a menos que reciba alimento de un compañero. Un sistema de camaradería asegura que la distribución de comida entre los murciélagos resulte equitativa*

Gerald S. Wilkinson

Por la noche, mucho después de que la mayoría de los depredadores visuales hayan dejado de merodear, los vampiros emergen de sus posaderos y emprenden un vuelo raso sobre el terreno, en busca de presas de sangre caliente. Pasadas una o dos horas, encontradas las víctimas apropiadas y habiéndose alimentado de su sangre, los murciélagos retornan a su escondrijo para dormir, alimentar a sus crías y relacionarse con sus compañeros.

Hasta hace poco se desconocía casi todo acerca del comportamiento y del ciclo biológico del vampiro común, *Desmodus rotundus*. Durante muchos años los biólogos estuvieron más interesados en la fisiología del animal que en su organización social, que se suponía bastante sencilla. Sin embargo, varios estudios recientes revelan que los vampiros son murciélagos muy sociales: las hembras se agrupan durante el día; por la noche, se distribuyen de nuevo entre la población, instaurando así una organización social fluida que se mantiene durante muchos años. Además, sabemos ya que las asociaciones duraderas entre las hembras permiten que los murciélagos regurgiten regularmente sangre, que aprovechan sus compañeros, con lo que aumentan de modo significativo sus probabilidades de supervivencia.

La razón para el comportamiento de regurgitación la revelaron los estudios realizados hace más de 15 años por Brian K. McNab, de la Universidad de Florida, quien demostró que el vampiro muere si no puede comer durante dos noches seguidas. Después de 60 horas sin comida pierde hasta el 25 por ciento de su peso, y ya no puede mantener una temperatura corporal crítica. Para abastecer de carburante al motor metabólico del cuerpo y evitar la muerte, estos murciélagos deben consumir el 50 (a veces incluso el 100) por ciento de su peso corporal en sangre cada noche.

Pero comer no siempre es fácil, en especial para los jóvenes, que deben aprender a morder rápidamente y hacerlo sin causar dolor a sus víctimas. He visto a caballos que erguían la cabeza, fustigaban con la cola y se frotaban contra cualquier objeto para desembarazarse de un vampiro hambriento. Aunque los murciélagos replican a estas estrategias defensivas volviendo al mismo animal (un blanco conocido) varias noches sucesivas, o bien alimentándose secuencialmente de heridas frescas, cada noche del 7 al 30 por ciento de los vampiros de cada grupo no consiguen obtener su ración de sangre. Al solicitar alimento a un compañero de reposadero, el murciélago puede eludir la muerte por inanición (al menos durante una noche más) y obtener así otra posibilidad de conseguir comida.

En 1978, Uwe Schmidt, zoólogo de la Universidad de Bonn, presentó las primeras pruebas de que las hembras regurgitan sangre a sus crías. Por aquel entonces Schmidt había criado murciélagos durante más de 10 años en una torrecilla del Poppelsdorfer Schloss, un viejo castillo que hoy es el principal edificio de investigación del instituto de zoología de la universidad, y había pasado gran parte de su carrera observando su comportamiento. Schmidt descubrió, por ejemplo, que, poco después del nacimiento, las madres suministran a las crías sangre regurgitada (además de la leche); encontró también que, en algunos casos, una cría tomaba sangre de un adulto que no era su madre. En un caso incluso observó una cría huérfana que era amamantada por una madre adoptiva. Compartir el alimento de esta manera, en la que algunos individuos participan en la alimentación de otros miembros del grupo no emparentados, es fenómeno rarísimo en los mamíferos; además de los murciélagos, se sabe que exhiben este comportamiento sólo unas pocas especies: perros salvajes, hienas, chimpancés y seres humanos.

Compartir la comida parece ser una actividad altruista: un murciélago donante ofrece comida (que de otro modo podría ser utilizada para su propia supervivencia o para la de su progenie) a un murciélago receptor, cuyas probabilidades de supervivencia se ven así aumentadas sin coste aparente para éste. Ahora bien, el altruismo auténtico no se había documentado nunca en animales que no pertenecieran a la especie humana, por la razón probable de que tal sistema unidireccional no resulte estable desde el punto de vista evolutivo.

GERALD S. WILKINSON ha pasado más de 10 años estudiando el comportamiento social de los murciélagos. Su interés arranca de un curso de verano realizado en Costa Rica, en 1978. Desde entonces ha investigado esos quirópteros en Sudamérica, Africa, Australia, el Sudeste asiático y los Estados Unidos. Wilkinson inició su formación en la Universidad de California en Davis, doctorándose por San Diego en 1984. Da clases de zoología en la Universidad de Maryland en College Park desde 1987.

**1. VAMPIRO COMUN, *Desmodus rotundus*. Se encuentra desde México hasta Argentina y Chile; prefiere las regiones donde las tierras han sido convertidas en pastos. Las hembras, véanse las de la fotografía adjunta, se reúnen en cuevas y árboles vacíos durante el día, saliendo sólo de noche en busca de presas de sangre caliente.**

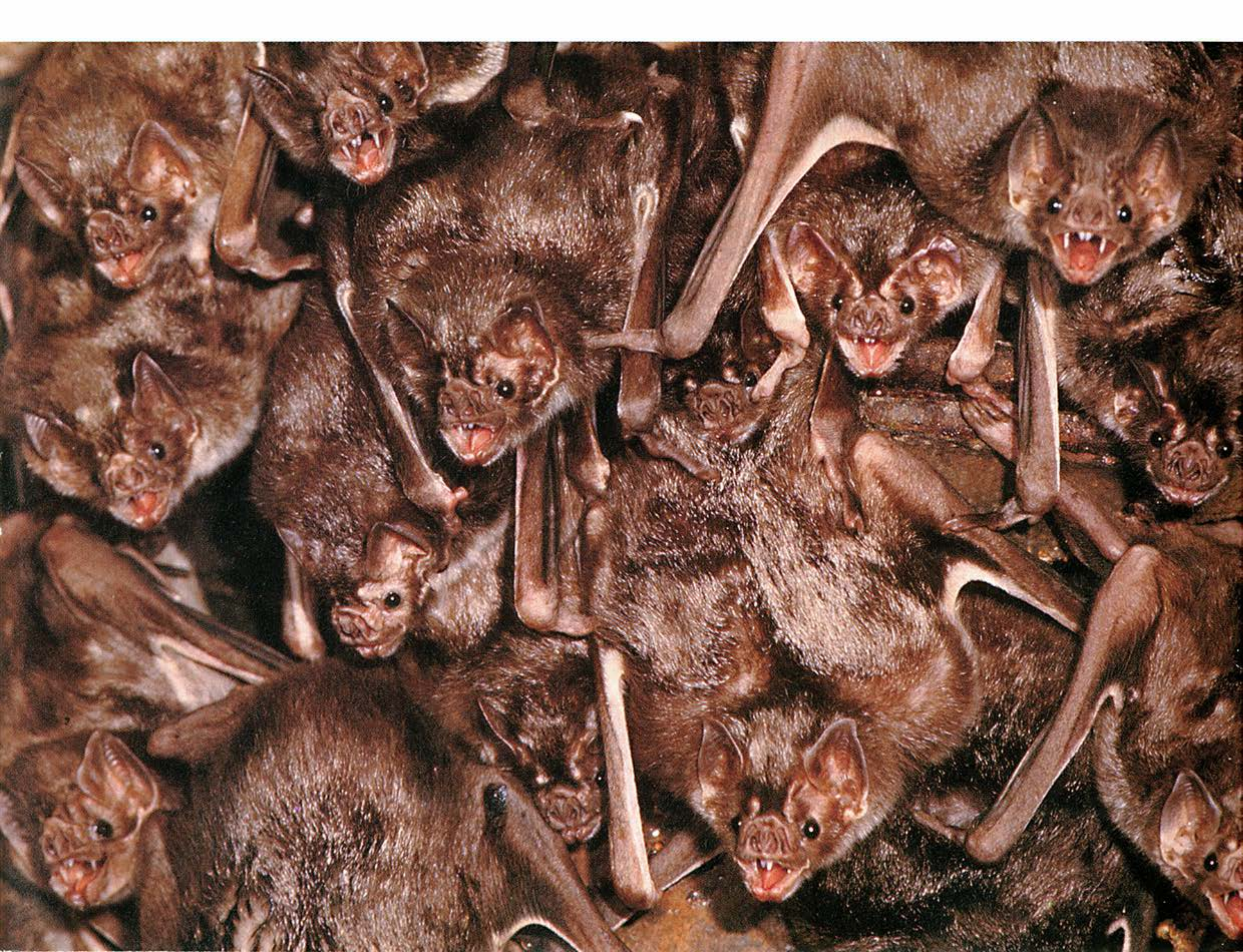
Ello se debe a que los individuos donantes, que ceden recursos, pierden a la larga la capacidad de competencia con los individuos receptores, que tendrán más recursos y, por ello, sobrevivirán a los primeros dejando más descendientes y transmitiendo así una mayor proporción de sus genes a la siguiente generación. Estudios minuciosos de comportamiento altruista realizados por varios investigadores revelan que muchos actos de altruismo aparente ocurren, en realidad, entre parientes (y constituyen, por tanto, una forma de selección de parentesco) o bien entre individuos que intercambian recursos sobre una base más o menos equitativa, en cuyo caso puede considerarse que son actos de altruismo recíproco, o de reciprocidad.

Enterado de las investigaciones de Schmidt, me propuse estudiar a los vampiros en su hábitat natural, para ver si la regurgitación de sangre era un acto de selección de parentesco o de reciprocidad (o de ambas cosas a la vez). Me dirigí a Costa Rica y allí, con la ayuda de mis colaboradores Robin Weiss, Michael L. Jones y Terri Lamp, estudié una población de *Desmodus rotundus* durante 26 meses, entre 1978 y 1983.

Contaba con averiguar, mediante la observación de su comportamiento de regurgitación, si los murciélagos alimentaban sólo a sus parientes, y por tanto se hallaban implicados en selección de parentesco, o si se intercambiaban alimento de forma recíproca e indistintamente con individuos emparentados y no emparentados, tratándose en dicho caso de un comportamiento de reciprocidad. Para confirmar la reciprocidad me faltaba por demostrar que se cumplían cinco condiciones: que las hembras se asociaban durante períodos prolongados, de modo que cada una de ellas tenía un número grande, aunque impredecible, de oportunidades para compartir sangre; que la probabilidad que presentaba un individuo de regurgitar a un determinado compañero de refugio podía predecirse a partir de su anterior relación; que los papeles de donante y receptor se invertían con frecuencia; que los beneficios a corto plazo para el receptor superaban los costos para el donante, y que los donantes eran capaces de reconocer a los tramposos y expulsarlos del sistema.

Los vampiros, que son comunes en América tropical allí donde la tierra ha sido convertida en pastos y hay ganado, son sujetos ideales para un estudio de esta naturaleza. Escogí como área de investigación un rancho ganadero en el noroeste de Costa Rica: Hacienda La Pacífica. (Desde entonces se transformó en estación de investigación ecológica, denominada hoy Centro Ecológico La Pacífica.)

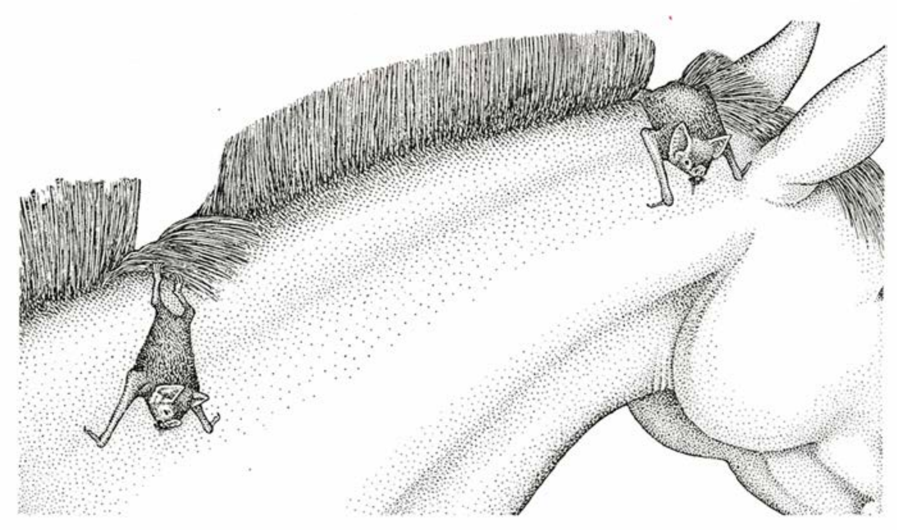
Allí descubrí que los vampiros, en ausencia de cuevas, pasan el día en árboles huecos, donde las temperaturas son constantes, la humedad elevada y se encuentran a oscuras incluso durante el día. La mayoría de los árboles de La Pacífica tenían una sola abertura en su base: nos introducíamos por la abertura y mirábamos hacia arriba con ayuda de binoculares y con una fuente de luz difusa, para observar a los murciélagos durante pe-





**2. RASGOS FACIALES de un vampiro; reflejan una dieta que tiene la sangre por componente básico. Las grandes orejas le ayudan a buscar la presa y a guiarse mediante ecolocación: el murciélago emite sonidos de alta frecuencia que son reflejados por los objetos que se encuentran en su trayectoria, y capta los ecos con las orejas. Receptores olfativos y térmicos situados en la nariz ancha y carnosa le permiten orientarse hacia la víctima adecuada. Una vez hallado un punto apropiado para morder, rasga ligeramente la carne con sus incisivos afilados como navajas. Después, con la ayuda de un anticoagulante de su saliva, lame la sangre que mana de la herida de su víctima.**

ríodos de varias horas. Aumentando gradualmente la cantidad de luz (a lo largo de un período de varios meses), acabamos por habituar a los murciélagos a nuestra presencia y pudimos observar sus interacciones. La entrada única en la base del árbol ofrecía otra ventaja: podíamos extender una fina red ante la misma y capturar a los murciélagos cuando salían de noche para ir a cazar. De este modo pudimos marcarlos con etiquetas y posteriormente cuantificar las pautas individuales de comportamiento.



**3. VAMPIROS COMUNES: se alimentan de diversos mamíferos, aunque prefieren, entre sus presas, los caballos y las mulas. Por lo general se posan sobre la cola o la crin del animal, que les proporciona un agarradero durante los veinte o treinta minutos que requieren para beber su ración de sangre. La víctima no coopera siempre: puede sacudir la cabeza y fustigar la cola en un intento de liberarse del murciélago. Un vampiro que no consigue alimentarse durante dos noches seguidas morirá de inanición, a menos que pueda solicitar y obtener sangre regurgitada de un compañero de escondrijo.**

Descubrimos que los murciélagos salían en busca de presas cada noche, a una hora variable con la fase de la luna: si había demasiada luz en el exterior, los vampiros esperaban hasta que la luna descendía. A diferencia de las otras dos especies de murciélagos chupadores de sangre (el vampiro de alas blancas, *Diaemus youngi*, y el vampiro de patas peludas, *Diphylla ecaudata*), que se alimentan sobre todo de sangre de aves, *Desmodus rotundus* se nutre en particular de sangre de mamíferos. El vampiro parece preferir los caballos a las vacas, y los localiza mediante una combinación de olfato, sonido y ecolocación.

Una vez ha identificado a las víctimas, el murciélago se posa por lo general sobre la cola o la crin del animal y cuelga de ella mientras busca un punto adecuado para morder. Células especializadas termosensoriales de la nariz guían al vampiro en su búsqueda de una región donde los vasos sanguíneos de la víctima se hallen cerca de la superficie. Cuando el vampiro encuentra ese punto idóneo, corta rápidamente un pequeño fragmento de piel con sus incisivos superiores, afilados como una navaja, y empieza a comer. Un anticoagulante de la saliva del vampiro permite que la sangre fluya durante los 20 o 30 minutos necesarios para consumir su ración. Después, el murciélago, con su estómago visiblemente henchido, vuelve al reposadero, donde permanece hasta la noche siguiente. A medida que los murciélagos retornaban, justo antes del amanecer, a los árboles que les servían de escondrijo, los atrapábamos con las redes para determinar qué individuos habían obtenido con éxito su ración de sangre.

Nuestro primer trabajo, al llegar a La Pacífica, fue el de marcar a todos los murciélagos de nuestra área de estudio con fajas ligeras de colores distintos; a las hembras se las ceñía una faja en la muñeca izquierda; a los machos, otra en la muñeca derecha. Cada faja tenía también pegada una tira de cinta reflectante, que nos permitía ver a los individuos posados incluso cuando los niveles de luz en el árbol eran muy bajos. Terminado el proceso de marcado, podíamos coger un murciélago al azar y registrar su comportamiento durante un período de 10 segundos, 100 veces consecutivas. Trabajando en parejas marcamos 600 vampiros y acumulamos más de 400 horas de observaciones de comportamiento. Una vez por semana censábamos los animales de cada árbol para documentar las pautas de





**4. ESTUDIAR LOS VAMPIROS** en su hábitat natural requiere pasar muchas horas en una posición incómoda (*izquierda*). Para observar a los murciélagos, que se guarecen en árboles huecos (así como en cuevas y en otros lugares oscuros), el autor y sus ayudantes pasaban de dos a seis horas diarias tendidos, mirando hacia arriba con una luz difusa. Bandas coloreadas colocadas alrededor de las muñecas de los vampiros (*derecha*) posibilitaron la labor de identificación de los individuos y el seguimiento de sus interacciones a lo largo de un período de cinco años (entre 1978 y 1983).

asociación entre los murciélagos que ocupan el mismo posadero diurno. También instalamos radiotransmisores en 37 individuos, para así determinar la extensión y el grado de superposición de sus áreas de búsqueda de alimento.

Descubrimos que la organización de los vampiros está dominada por grupos de ocho a 12 hembras adultas y un número igual de crías (una por cada hembra). Las crías nacen durante todo el año a intervalos de unos 10 meses; las hembras permanecen junto a su madre, mientras que los machos se van cuando alcanzan la madurez reproductora, lo que tiene lugar entre los 12 y los 18 meses de edad. Contrariamente a lo que ocurre en otras especies de murciélagos tropicales, en las que los machos defienden "harenes", los vampiros machos defienden territorios. Forman jerarquías de dominancia en el interior de árboles huecos, luchan entre sí para obtener la posición alfa, situada cerca de la parte más alta de la cavidad hueca del tronco (donde con frecuencia se agrupan las hembras), y defienden vigorosamente sus territorios, a veces hasta la muerte, contra los intrusos. Estos son machos que generalmente reposan solos o en grupos pequeños durante el día, en árboles que no suelen visitar las hembras.

En nuestro seguimiento de la conducta de los vampiros y anotación de sus asociaciones, nos sorprendió el carácter estable, y a la vez fluido, de su organización social. La población de vampiros en nuestra área de estudio podía subdividirse en tres grupos de unas 12 hembras adultas, cada uno de los cuales solía admitir agrupamientos menores. Aunque los tres grupos estaban aislados entre sí, la composición de cada subgrupo dentro de cada grupo variaba continuamente. Cada grupo tenía derechos exclusivos sobre un total de unos seis árboles, y una o dos veces por semana las hembras intercambiaban su refugio (llevándose en ocasiones consigo sus crías a otro árbol), y se redistribuían durante el proceso.

De las crías, las hembras permanecen con su madre tras la pubertad; por eso conviven varias generaciones agrupadas en un árbol. Pero mis análisis bioquímicos de muestras de sangre sugieren que sólo alrededor de un 50 por ciento de las crías de un grupo comparten el mismo padre. Presumo que ello se deba a que las hembras no presentan fidelidad a un árbol determinado (mientras que los machos sí), con lo que se hallan periódicamente expuestas a nuevos machos, con los que a veces se aparean. Además, por razones que no se conocen bien, las hembras cambian de grupo de vez en cuando (tal vez porque las presas se han hecho difíciles de localizar); por término medio, una nueva hembra se incorpora a un grupo cada dos años. En consecuencia, cada grupo está constituido por varias matrilíneas, dentro de las cuales el grado de parentesco es alto, pero entre ellas es bajo.

El análisis de las asociaciones de reposadero revela que las hembras adultas exhiben una preferencia por determinadas hembras, que no puede explicarse simplemente sobre la base de alguna característica física del lugar de reposadero. Además, parece que su preferencia no es siempre por parientes, sino que puede ser por otras hembras no emparentadas, hallazgo que es coherente con la teoría de la reciprocidad. Una vez hubimos determinado que los murciélagos poseían una afinidad mutua, necesitábamos contestar a la siguiente pregunta: ¿permanecen los vampiros juntos durante períodos prolongados y tienen así la oportunidad de desarrollar y mantener un sistema de ayuda mutua?

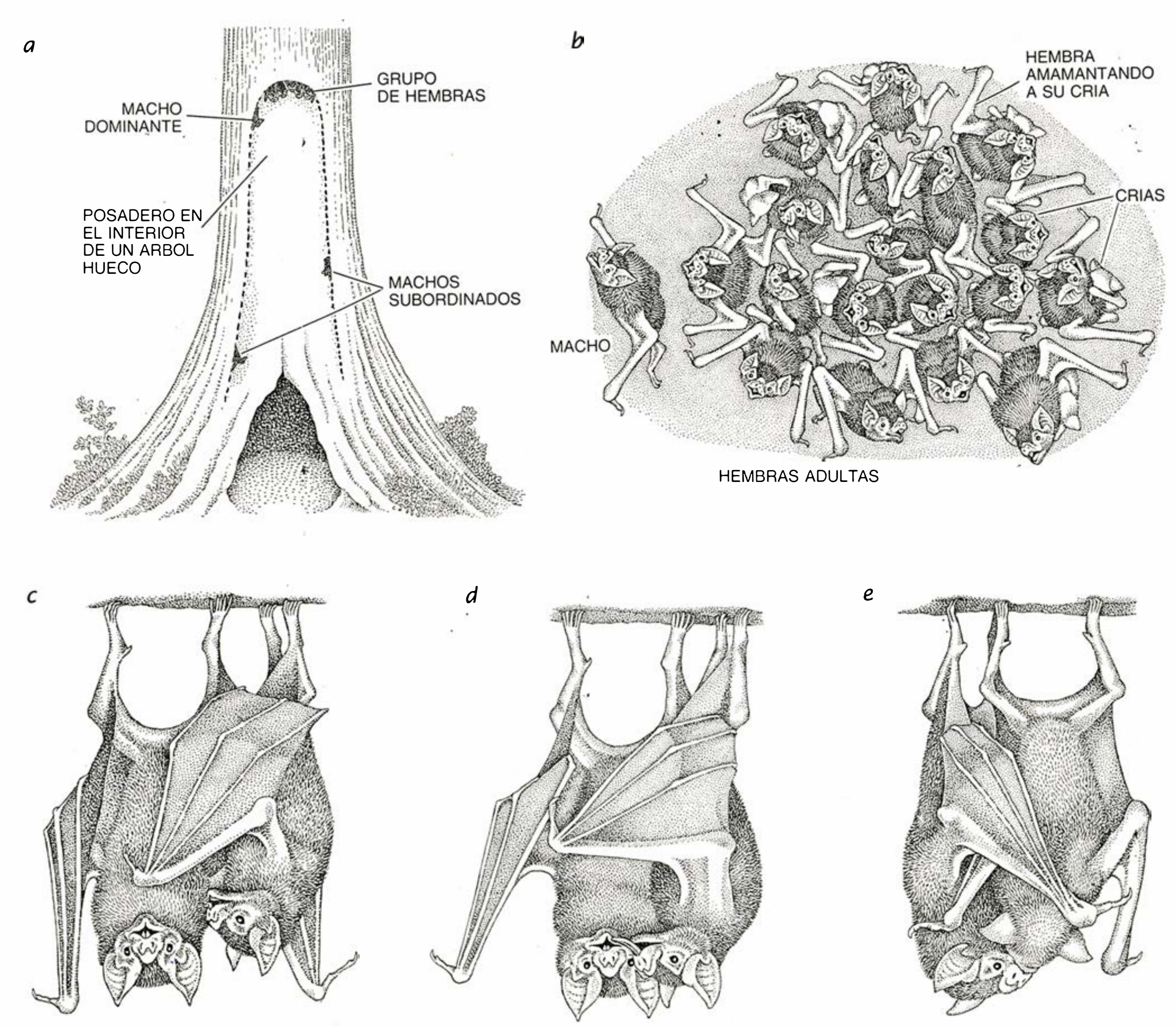
La respuesta parece ser afirmativa. Rexford D. Lord, hoy en la Universidad de Indiana, determinó la esperanza de vida máxima de los vampiros contando los anillos de crecimiento anual de sus dientes; descubrió que las hembras pueden vivir hasta 18 años. Y a partir de estudios de anillamiento efectuados en La Pacífica por Theodore H. Fleming, de la Universidad de Miami, en la década de los setenta, sabíamos que al menos dos de las hembras de nuestra área de estudio habían reposado juntas durante más de 12 años. A la vista de su longevidad y del hecho de que en una población siempre había individuos que no conseguían comer periódicamente, llegamos a la conclusión de que los vampiros cumplían el primer requisito de reciprocidad: los individuos no sólo creaban relaciones mutuas prolongadas y estables, sino que siempre existía la oportunidad de compartir la comida.

El siguiente paso consistió en determinar si la sangre se regurgitaba al azar dentro de un grupo o bien si las

hembras regurgitaban sólo a parientes próximos o a compañeros previos de reposadero, como predecían, respectivamente, la selección de parentesco y la teoría de la reciprocidad. Para ello, precisábamos estimar la frecuencia relativa con que cada individuo compartía la sangre en condiciones naturales.

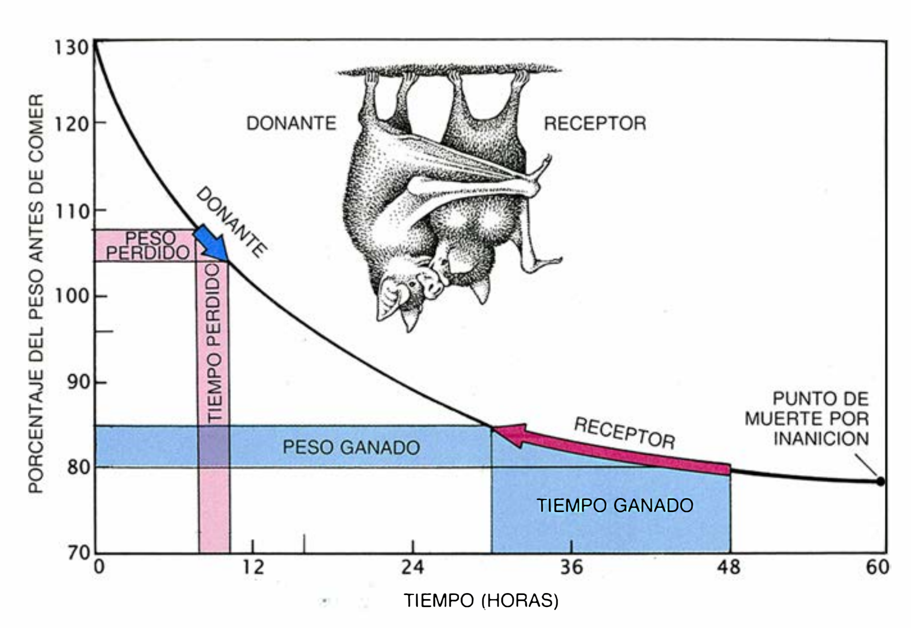
A lo largo de los cinco años de nuestro estudio fuimos testigos de un total de 110 casos en los que los vampiros regurgitaron sangre para compartirla. El 70 por ciento de las regurgitaciones tuvieron lugar entre una madre y su cría, por lo que podían considerarse cuidados maternos. Sin embargo, el 30 por ciento restante implicó a hembras adultas que alimentaban a crías que no eran las propias, hembras adultas que alimentaban a otras hembras adultas y, en dos ocasiones, machos adultos que alimentaban a su progenie.

¿Regurgitaban selectivamente o no los murciélagos? Comparamos el grado de parentesco entre un donante y un receptor, y cotejamos también su índice de asociación de posadero (proporción de veces que aparecían juntos los dos mismos individuos en la misma agrupación), con el propósito de establecer —sobre una u otra base— si los valores de regurgitación eran superiores entonces a los registrados cuando el receptor solicitaba aleatoriamente su parte de cualquier donante potencial del posadero. Descubrimos, en efecto, que ambos factores, el parentesco y la asociación previa en el escondrijo, constituían predictores importantes de la respuesta de un individuo a una solicitación de regurgitación. Nuestros resultados muestran que los vampiros no comparten el alimento aleatoriamente, sino que lo comparten de preferencia con individuos que suelen ser compañeros de reposadero y que frecuentemente, pero no siempre, están emparentados, hallazgo que concuerda a la vez con la teoría de la reciprocidad y con la de selección de parentesco.




**5. MACHOS Y HEMBRAS de vampiro se guarecen frecuentemente en el mismo árbol (*a*). Las hembras se arraciman cerca de la parte superior de la cavidad, a unos 3,5 metros o más del suelo, bajo la custodia de un macho dominante. Dos o tres machos subordinados ocupan el mismo árbol, pero se cuelgan en posiciones más cercanas al suelo. Hasta 12 hembras, cada una con una cría (los murceguillos difieren en tamaño porque los nacimientos tienen lugar a lo largo de todo el año), se reúnen en un mismo árbol (*b*). Aunque la composición de los grupos de reposo varía de un día para otro, algunas hembras se asocian durante muchos años y se regurgitan sangre entre sí, comportamiento que constituye una forma de altruismo recíproco. Primero un vampiro hambriento solicita sangre regurgitada de una compañera de reposadero acicalándola (*c*), es decir, lamiendo a la donante potencial bajo el ala, y luego lamiendo sus labios (*d*). Si la hembra donante es receptiva, responde regurgitando sangre (*e*). Unicamente los murciélagos que son parientes cercanos o que han mantenido, con anterioridad, una asociación prolongada comparten entre sí la sangre.**

El siguiente paso de nuestro estudio fue comprobar experimentalmente la reciprocidad. Si los vampiros practicaban el altruismo recíproco, podía entonces predecirse que los individuos ayudarían sólo a los expuestos en peligro inminente de inanición, y compensarían de preferencia a los murciélagos que, en ocasiones anteriores, los hubieran alimentado. Para comprobar estas hipótesis, capturamos a cuatro hembras adultas de nuestra principal área de estudio en La Pacífica y cuatro de una zona de estudio secundaria en el Parque Nacional de Santa Rosa, a unos 50 kilómetros al norte. Sabíamos por sus marbetes que dos de los murciélagos de La Pacífica eran abuela y nieta (que compartían por tanto un cuarto de su herencia genética); las otras no estaban emparentadas, pero presentaban un grado elevado de asociación de reposadero.




**6. ANALISIS DE COSTE-BENEFICIO que implica el compartir sangre entre los vampiros. A tenor del mismo, los receptores se benefician más de lo que pierden los donantes. El autor pesó hembras adultas que retornaban a la guarida después de comer; las pesó luego cada hora, durante las 24 horas siguientes. Un individuo que ha comido puede volver con el 130 por ciento de su peso de precomida (la mitad del peso de una ración de sangre se pierde con la orina durante la primera hora siguiente a la ingestión), mientras que un murciélago que no ha conseguido abastecerse durante dos noches seguidas podría retornar con el 80 por ciento de su peso anterior. Si regurgita cinco mililitros de sangre condensada a un compañero hambriento, el vampiro donante puede pasar del 108 al 103 por ciento de su peso de precomida, pero pierde únicamente tres horas del tiempo que le queda hasta la muerte por inanición. En cambio, el vampiro receptor gana 18 horas, con lo que su beneficio es mucho mayor de lo que pierde el donante. (El esquema ha sido realizado por Patricia J. Wynne.)**

Iniciamos nuestro experimento habituando a los murciélagos a la cautividad y también a ser alimentados cada noche con botellas de cristal graduadas, lo que nos permitía registrar la cantidad de sangre que cada murciélago ingería en una comida. Una vez se encontraron cómodos en sus jaulas, seleccionamos al azar uno por noche y lo encerramos en una jaula separada, en la que se le privaba de alimento. A la mañana siguiente devolvíamos el vampiro experimental a su jaula original y observábamos sus interacciones con sus compañeros de encierro.

Nuestros resultados revelaron que el fenómeno de compartir la sangre se producía casi siempre entre murciélagos de la misma población; sólo una vez se dio entre extraños. Además, descubrimos que el compartir la sangre no ocurría al azar, ni siquiera entre murciélagos con un grado elevado de asociación anterior. Sí parecía, por contra, que los murciélagos no emparentados desarrollaban un sistema de camaradería, de manera que dos individuos determinados se regurgitaban mutua y casi exclusivamente; prueba clara de que sus papeles se invertían de forma regular.

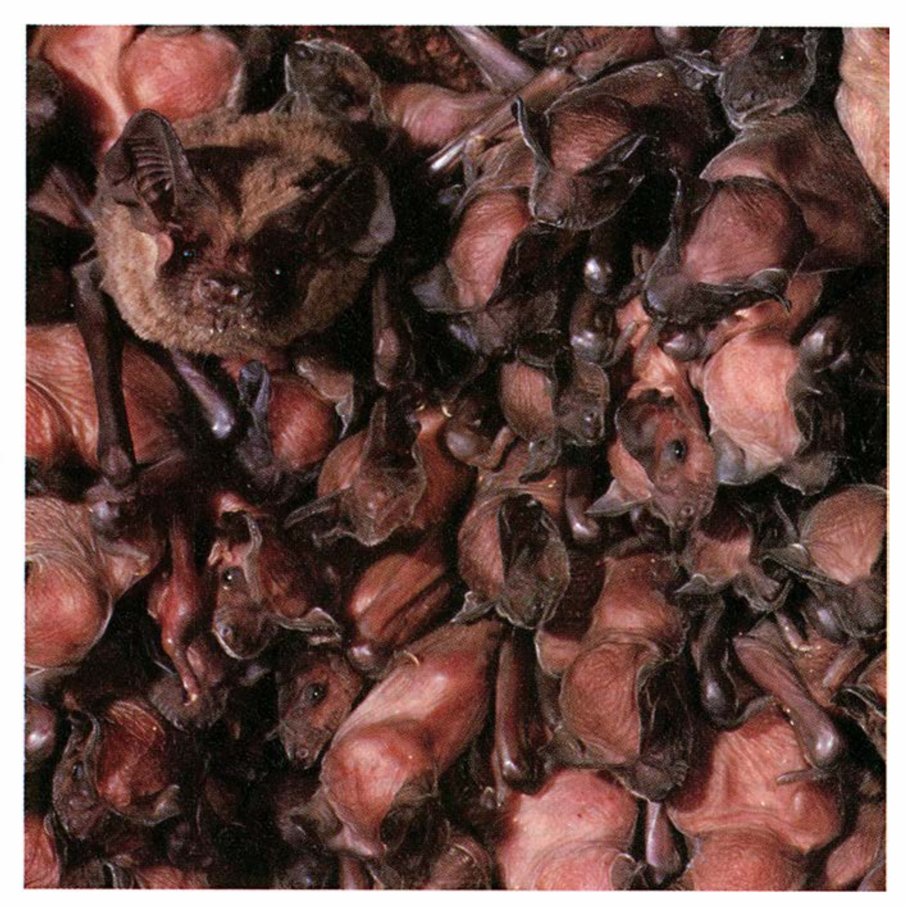
Otro criterio de la teoría de reciprocidad remite al coste de donar sangre; éste debe ser pequeño en relación al beneficio que proporciona al receptor. Dicho de otro modo, al regurgitar sangre a un compañero de reposadero, el donante debe ser capaz de salvar la vida de aquél sin arriesgar la suya propia. Para probar si esto era o no cierto en el caso de los vampiros, precisábamos medir los costos y los beneficios de compartir la sangre, y hacerlo de dos maneras: por vía directa, determinando la cantidad de sangre y la frecuencia de ingestión de la misma que se requiere para evitar la inanición; y por vía indirecta, estimando el efecto de compartir la sangre en la supervivencia a largo plazo, mediante simulaciones con ordenador.

Por el trabajo de McNab y el nuestro propio, sabíamos que un murciélago debía consumir de 20 a 30 mililitros de sangre cada 60 horas para evitar morirse de hambre. Sabíamos también que un murciélago al borde de la inanición ganaba hasta 12 horas de vida, y contaba así con otra oportunidad para encontrar comida, si un compañero de reposadero le suministraba sangre. Un compañero cooperativo, que haya comido recientemente y tenga, por tanto, una elevada tasa metabólica, pierde menos de 12 horas si dona una ración de sangre, de modo que todavía le quedarán 36 horas y dos noches de caza antes de alcanzar el punto de inanición. A tenor, pues, de la teoría de la reciprocidad, sólo los vampiros a los que les quedaran menos de 24 horas de vida debían ser alimentados por sus compañeros. Nuestro experimento con ejemplares cautivos, en el que les negábamos comida durante períodos de 24 horas, lo confirmó.

Para que un sistema de intercambio recíproco funcione, es necesario que los murciélagos se reconozcan entre sí y distingan, y excluyan, a los tramposos. Aunque todavía no hemos demostrado que haya tramposos, o que queden excluidos del sistema, tenemos razones para creer que los vampiros están capacitados para su reconocimiento individual. Digamos de entrada que hemos observado, en murciélagos cautivos, que sólo los que han coincidido en una asociación anterior se regurgitarán sangre uno a otro. Parece, por tanto, probable que se identifiquen entre sí.

Ciertas pruebas circunstanciales abonan la importancia del aseo social en el reconocimiento de los compañeros de reposadero. Los murciélagos pasan más del 5 por ciento de cada día acicalándose y lamiéndose unos a otros, tareas que parecen constituir un buen preludio para la regurgitación: los vampiros hambrientos asean con frecuencia a donantes potenciales (hembras que han comido recientemente). Al igual que ocurre con la regurgitación de sangre, el acicalamiento se da entre los individuos que son parientes cercanos, o bien que se han asociado previamente, con más frecuencia que entre murciélagos que son extraños.

Schmidt ha proporcionado pruebas adicionales de que los murciélagos se



**7. MURCIELAGOS RABUDOS MEXICANOS; mantienen a sus crías en jardines de infancia comunales, que pueden albergar un millón o más de murceguillos recién nacidos. A pesar de la densidad de crías (hasta 40 en un espacio de 100 centímetros cuadrados), las madres, como la que hay en la parte superior izquierda, localizan a sus hijos el 83 por ciento de las veces, de manera que sólo muy raramente alimentan a crías que les son extrañas. (Fotografía obtenida por Merlin D. Tuttle.)**

reconocen entre sí. Mediante el análisis de sonogramas, Schmidt y sus alumnos han descubierto que los murciélagos emiten vocalizaciones que difieren de un sujeto a otro. Estas "llamadas de contacto" acompañan con frecuencia a las sesiones de aseo y tienen las características acústicas (frecuencia variable e intensidad baja) que se precisan para codificar la identidad del individuo. Señales auditivas (y posiblemente señales olfativas) de este tipo permiten probablemente reconocer a los camaradas y a los tramposos que solicitan sangre y en cambio ellos nunca la ofrecen.

Parece, pues, que tanto la reciprocidad como el reconocimiento de parentesco promueven el que los compañeros de posadero compartan sangre, y que la regurgitación sea beneficiosa (al menos a corto plazo) para un receptor. ¿De qué modo este comportamiento, tan costoso desde el punto de vista energético, puede afectar a las tasas de supervivencia globales dentro de una población? Para establecer que la reciprocidad persiste en un sentido evolutivo, hemos de demostrar, al menos en teoría, que los vampiros que comparten comida con sus compañeros poseen una tasa de supervivencia anual superior a la de los que no lo hacen.

A partir de nuestros estudios de captura con redes sabíamos que, por término medio, alrededor del 30 por ciento de los vampiros inmaduros (los de menos de dos años) no consiguen obtener su ración de sangre en una noche dada, mientras que sólo fracasan el 7 por ciento de los machos y hembras reproductivamente maduros. Las observaciones de campo nos habían enseñado también que el fracaso parece ser aleatorio: todos los individuos de un determinado grupo de edad tienen la misma probabilidad de volver de vacío. Con esta información y la ayuda de un ordenador determinamos que la mortalidad anual para los adultos no socorridos por otros (considerando que el 7 por ciento de los adultos no consiguen alimentarse en una noche determinada, y que dos noches seguidas sin sangre conducen a la muerte) debe ser del orden del 82 por ciento. Puesto que la mortalidad en los vampiros adultos es de sólo el 24 por ciento anual, llegamos a la conclusión de que compartir el alimento viene primado por la selección natural.

Los vampiros han desarrollado un sistema de intercambio de alimento por el que comparten sangre con sus compañeros menesterosos. Aunque este comportamiento supone un riesgo para el donante, el receptor tiene más probabilidades de sobrevivir otra noche. Además, nuestros estudios han demostrado que los individuos que regurgitan sangre a sus compañeros de posadero obtienen una ventaja inmediata, que se traduce en el aumento de su propia supervivencia y en ocasiones de la supervivencia de sus parientes. Ambas, pues, reciprocidad y selección de parentesco, operan en este sistema.

¿Es el compartir el alimento un fenómeno exclusivo de los vampiros? No. Hembras de varias especies de murciélagos insectívoros, entre ellas el murciélago troglodita, *Miniopterus schreibersi*, el murciélago rabudo mexicano, *Tadarida mexicana*, y el murciélago vespertino, *Nycticeius humeralis*, alimentan a crías distintas de las suyas.

El murciélago troglodita no ha sido objeto todavía de estudio detallado. Gary F. McCracken y sus alumnos, de la Universidad de Tennessee en Knoxville, han venido dedicándose a las colonias de cría de murciélagos rabudos durante los últimos nueve años. Varios millones de murciélagos rabudos se congregan en cuevas de los Estados Unidos sudoccidentales cada verano, para parir sincrónicamente a sus crías. Los juveniles son mantenidos en jardines de infancia, donde la densidad es muy alta, de hasta 40 crías por 100 centímetros cuadrados. Las hembras descansan en un lugar apartado y visitan a sus crías sólo dos veces al día para alimentarlas. Cuando las hembras se acercan, las crías hambrientas se agolpan a su alrededor; se han observado hasta cuatro crías que intentaban comer secuencialmente de una misma hembra.

Cualquier observador diría que las hembras alimentan a las crías que primero se les acercan. Por esta razón muchos investigadores pensaron que las hembras constituían un recurso común; después de todo, encontrar la propia cría entre millones de otras parecía imposible. Pero McCracken ha demostrado, comparando las variaciones en las enzimas sanguíneas entre las hembras lactantes y los murciélagos mamantones, que las madres encuentran y amamantan a sus propias crías el 83 por ciento del tiempo, guiadas, tal parece, por señales olfativas y auditivas. Los amamantamientos no maternofiliales obedecen se-

guramente a robos de leche por parte de crías agresivas.

En especies que forman semejantes agregaciones, como el murciélago rabudo, los beneficios de los jardines de infancia (las crías conservan mejor el calor y se reduce el riesgo de depredación de cada individuo) sobrepasan el coste de alimentar ocasionalmente una cría extraña. En los murciélagos rabudos, el amamantamiento no maternofilial (cuando tiene lugar) no parece ser una forma de selección de parentesco ni de altruismo recíproco, sino fruto de un error aleatorio.

He pasado los dos últimos veranos investigando con mis alumnos el amamantamiento no maternofilial en los murciélagos vespertinos del norte de Missouri. Estos, como los rabudos, forman colonias de cría. Se trata, empero, de colonias bastante pequeñas (alojan sólo de 30 a 200 adultos) y, por lo general, se esconden en buhardillas y no en cuevas. A diferencia del murciélago rabudo, que pare una cría cada verano, el murciélago vespertino pare generalmente dos y, en ocasiones, tres crías cada vez. Nuestros estudios indican que una madre cría fielmente a sus hijos durante la primera semana de vida, pero que a medida que los jóvenes murciélagos crecen tienden, de vez en cuando, a obtener alimento de otras hembras. En torno a las tres semanas de edad, la cría suele recibir alimento de una hembra distinta de su madre en un 20 por ciento de las ocasiones.

¿Se trata de un caso de altruismo recíproco o de selección de parentesco? Puesto que puede observarse a las hembras aceptando o rechazando activamente a las crías que les solicitan leche, parece que están discriminando de algún modo entre los murceguillos. Los resultados preliminares apoyan la idea de una alimentación selectiva de las hembras a sus parientes: análisis de marcadores enzimáticos sanguíneos indican que, con gran frecuencia, las hembras no maternas son primas de las crías a las que amamantan. Los datos recogidos por mi alumno de doctorado J. Andrew Scherrer señalan, además, que cada murceguillo tiene una llamada única y que las realizadas por miembros de una misma familia comparten rasgos comunes. Quizá las hembras no maternas pueden reconocer a las crías emparentadas evaluando el grado de similaridad entre sus llamadas con las de su propia cría.

La investigación sobre el fenómeno de compartir la comida en los murciélagos ilustra un tema habitual en biología evolutiva: comportamientos similares que se exhiben por especies distintas pueden proceder de presiones evolutivas enteramente diferentes. Aunque se acepta que la selección de parentesco constituye una fuerza evolutiva poderosa y universal, bajo determinadas circunstancias (pensemos en los casos en que los animales viven en pequeños grupos y se da una alta probabilidad de prestarse mutuo socorro) adquiere mayor verosimilitud que la reciprocidad resulte más beneficiosa que la selección de parentesco, siempre y cuando los tramposos puedan ser detectados y excluidos del sistema.

Una mayor comprensión de las fuerzas responsables de la evolución social en los vertebrados requiere la identificación de los mecanismos que subyacen bajo el reconocimiento de parentesco y el reconocimiento individual. Nuestra investigación sobre el comportamiento de ayuda mutua entre murciélagos demuestra que el papel que la selección de parentesco y la reciprocidad desempeñan en una sociedad depende de la capacidad de reconocimiento de los animales de esa sociedad.

La determinación del grado en que los individuos se reconocen e interactúan de preferencia con los parientes recibirá un impulso vigoroso con las técnicas moleculares modernas (como la caracterización fina del ADN), lo que aumenta de manera significativa la capacidad de los investigadores a la hora de medir el parentesco entre los animales en el campo. Debido a que los murciélagos poseen un sistema auditivo complejo que les permite guiarse y capturar presas, creo que el estudio minucioso de sus vocalizaciones en los distintos contextos de su vida social puede proporcionar una apasionante información sobre los mecanismos mediante los cuales los animales reconocen a sus parientes y a sus asociados inmediatos. A su vez, los resultados de tal estudio habrán de aportar abundante luz sobre el comportamiento social de los vertebrados en general.


BIBLIOGRAFIA COMPLEMENTARIA

THE EVOLUTION OF COOPERATION. Robert Axelrod. Basic Books, Inc., 1984

SOCIAL EVOLUTION. Robert Trivers. Benjamin-Cummings Publishing Co., 1985.

THE NATURAL HISTORY OF VAMPIRE BATS. Dirigido por Arthur M. Greenhall y Uwe Schmidt. CRC Press, 1988.

RECIPROCAL ALTRUISM IN BATS AND OTHER MAMMALS. Gerald S. Wilkinson en *Ethology and Sociobiology*, vol. 9, n.os 2-4, págs. 85-100; julio de 1988.

# Transiciones de fase en las perovskitas

*La sencillez estructural de estos compuestos ha permitido establecer relaciones entre sus variaciones de simetría, transiciones de fase y propiedades físicas*

Juan Bartolomé Sanjoaquín y Rafael Navarro Linares

Las perovskitas han suscitado interés en dos campos distintos, aunque complementarios. En ciencia básica, estos materiales han tenido carácter pionero en el descubrimiento y comprensión de muchos fenómenos del estado sólido, como la ferroelectricidad, asociándolos a cambios en los estados o posiciones de los átomos constituyentes. En el dominio de la tecnología, algunos materiales con estructura de perovskita son utilizados como sensores, componentes electrónicos, abrasivos y otros, de obvio interés económico. De hecho, con las perovskitas se ha producido un proceso de realimentación, por el cual las aplicaciones técnicas han incentivado la ciencia fundamental y viceversa.

Con la aparición de los superconductores de alta temperatura crítica se ha renovado el interés en las perovskitas. Materiales como $La_{2-x}Sr_xCuO_{4-\delta}$ o $YBa_2Cu_3O_{7-\delta}$, en los que se produjeron los primeros hallazgos, presentan estructuras derivadas de tales compuestos [véase "Perovskitas", por R. M. Hazen; INVESTIGACIÓN Y CIENCIA, agosto de 1988]. En estos materiales se evidencia el paradigma ciencia-tecnología; existe un interés científico básico, ya que no se conoce bien el mecanismo que produce la superconducción, y un interés tecnológico, porque sus propiedades pueden revolucionar la industria eléctrica y electrónica. En efecto, desde hace cuatro años se vive una auténtica carrera mundial en la obtención de superconductores con mayor temperatura crítica (superior a 200 grados Kelvin en el $Y_5Ba_6Cu_{11}O_4$ cuando se publica este artículo), y con corrientes críticas competitivas similares a los materiales convencionales, así como en la búsqueda de posibles aplicaciones. Con nuestros antecedentes de trabajo en perovskitas, ha resultado natural participar, en colaboración con otros laboratorios nacionales y extranjeros, en el estudio de los nuevos superconductores.

En este artículo pretendemos ilustrar las propiedades más relevantes de las perovskitas, mostrar los cambios de fase más característicos, la variación de propiedades que producen y las herramientas experimentales que permiten su estudio. Además, se avanzan, a la luz de los conocimientos actuales, las causas microscópicas de los comportamientos observados.

JUAN BARTOLOME SANJOAQUIN y RAFAEL NAVARRO LINARES trabajan en el Instituto de Ciencia de Materiales de Aragón (ICMA), centro mixto del Consejo Superior de Investigaciones Científicas y de la Universidad de Zaragoza. Ambos se formaron y obtuvieron su doctorado en ciencias físicas en dicha universidad y realizaron una estancia de dos años en el laboratorio Kamerling Onnes de Leiden. Bartolomé es profesor de investigación y director del ICMA y Navarro es catedrático.

## ¿Qué son las perovskitas?

Se da el nombre de perovskitas a varias familias de materiales naturales y sintéticos, de estructura semejante a la del mineral perovskita $CaTiO_3$. En una perovskita cúbica ideal $ABX_3$, los átomos A [*véase la figura 1a*] se sitúan en los vértices del cubo, los B en su centro y los X en el centro de las caras. La estructura está formada por grupos octaédricos, $BX_6$, muy estables, lo que permite una descripción poliédrica de la red cristalográfica como un conjunto de octaedros enlazados por sus vértices, donde los átomos A ocupan el centro de los espacios intermedios [*véase la figura 1b*]. La variedad posible de componentes básicos de las perovskitas, A, B, X, hace que el material pueda estar agregado por interacciones de distinta intensidad y carácter, dando lugar a propiedades diferentes según sea su constitución cristalográfica, magnética o eléctrica, es decir, sus fases. Así, cuando X es un anión como flúor ($F^-$), cloro ($Cl^-$) o bromo ($Br^-$), y A y B, cationes, da lugar a compuestos iónicos que serán buenos aislantes eléctricos. Sin embargo, se pueden obtener perovskitas con enlace esencialmente metálico, y por ello conductores eléctricos, como en los compuestos intermetálicos de manganeso (Mn) con nitrógeno, $Mn_3MN$, y carbono, $Mn_3MC$, donde M es un metal de transición.

Ningún científico cuestiona que el conocimiento completo de las interacciones elementales entre los componentes de la materia permite predecir y caracterizar nuevas fases, analizar procesos de cambio y comprender los mecanismos que los producen. Sin embargo, todo investigador sabe igualmente que esta tarea, objeto de la física de estado sólido, no es posible sin simplificaciones, y una de las mayores ayudas son el orden periódico inherente al estado cristalino y la simetría. Otra simplificación importante es la relativa a la localización de los electrones; por eso, consideraremos en este artículo, por orden, las perovskitas aislantes (con electrones localizados), las conductoras eléctricas (con electrones deslocalizados), los óxidos con estructura de perovskita, cuyos representantes cubren ambos extremos, y, finalmente, los nuevos óxidos superconductores.

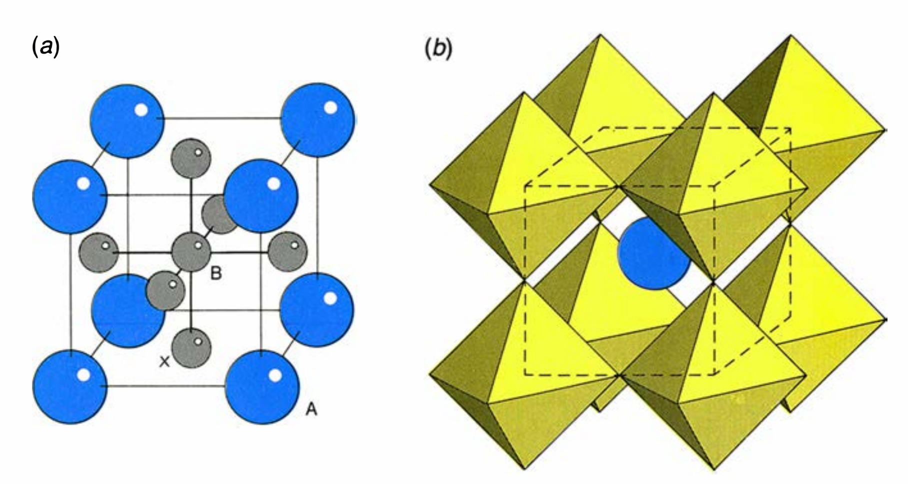
## Aislantes eléctricos

Empecemos por las perovskitas en cuanto aislantes eléctricos. En cualquier material, uno de los mecanis-

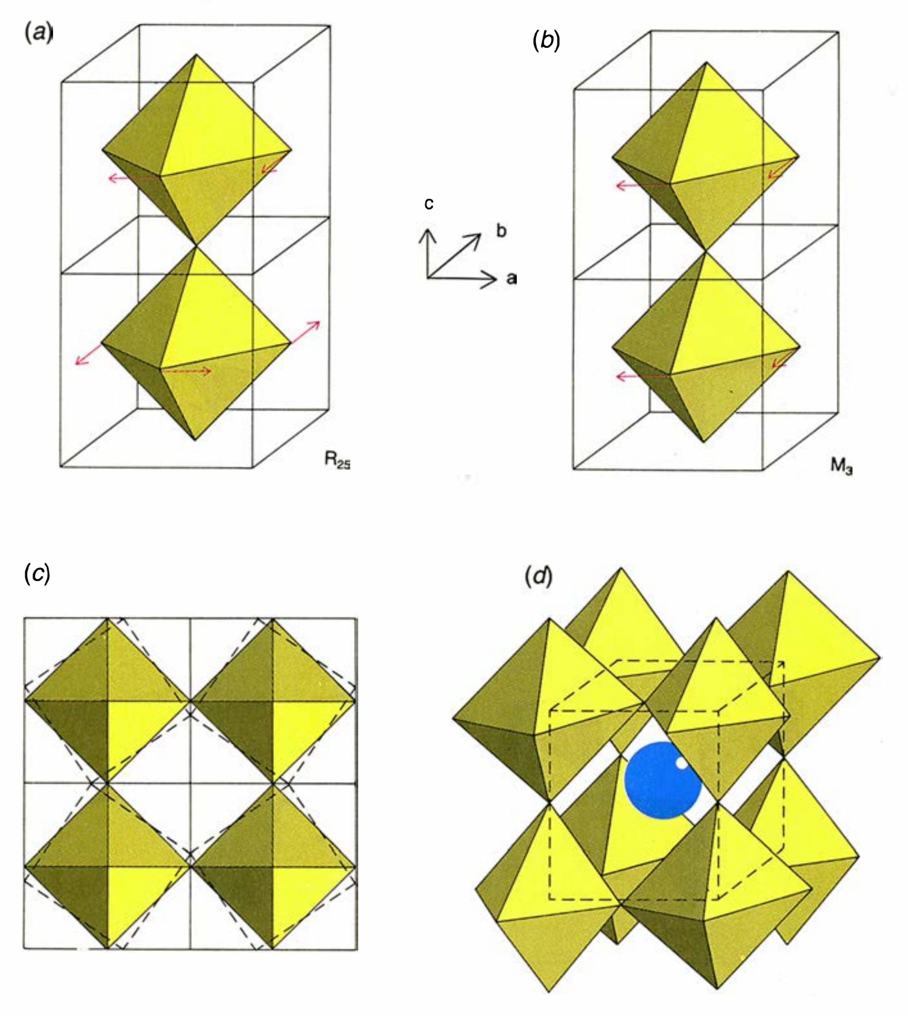
mos fundamentales de absorción de energía es la variación del estado de movimiento de los átomos que lo constituyen. La mayor intensidad de enlace B-X de los octaedros $BX_6$ facilita el estudio de las vibraciones atómicas colectivas de las perovskitas (*modos fonónicos*), porque permite su clasificación en modos internos de octaedros aislados y en movimientos relativos de los octaedros y de los átomos de la subred A.

La mayoría de las perovskitas cúbicas, al bajar la temperatura, sufren deformaciones que pueden describirse como giros de los octaedros alrededor de sus diagonales o como distorsiones de su regularidad. Más allá de la descripción de las estructuras deformadas, nos interesa su evolución, lo que exige conocer las posiciones atómicas y sus oscilaciones. En las perovskitas los modos fonónicos $M_3$ y $R_{25}$ adquieren especial interés, razón por la cual los describimos a continuación. Si, en una capa de octaedros [*véase la figura 2*], uno de ellos gira hacia la izquierda alrededor de la diagonal perpendicular a la capa, al compartir vértices los cuatro adyacentes giran hacia la derecha y originan un movimiento de vaivén coherente. A su vez, los octaedros de las capas inmediatas superior e inferior pueden oscilar a la par (*modo $M_3$*) o en oposición (*modo $R_{25}$*). Estos movimientos coordinados se llaman "modos normales" de la red y cualquier movimiento de los átomos que la constituyen pueden describirse como la superposición de todos los modos normales posibles.

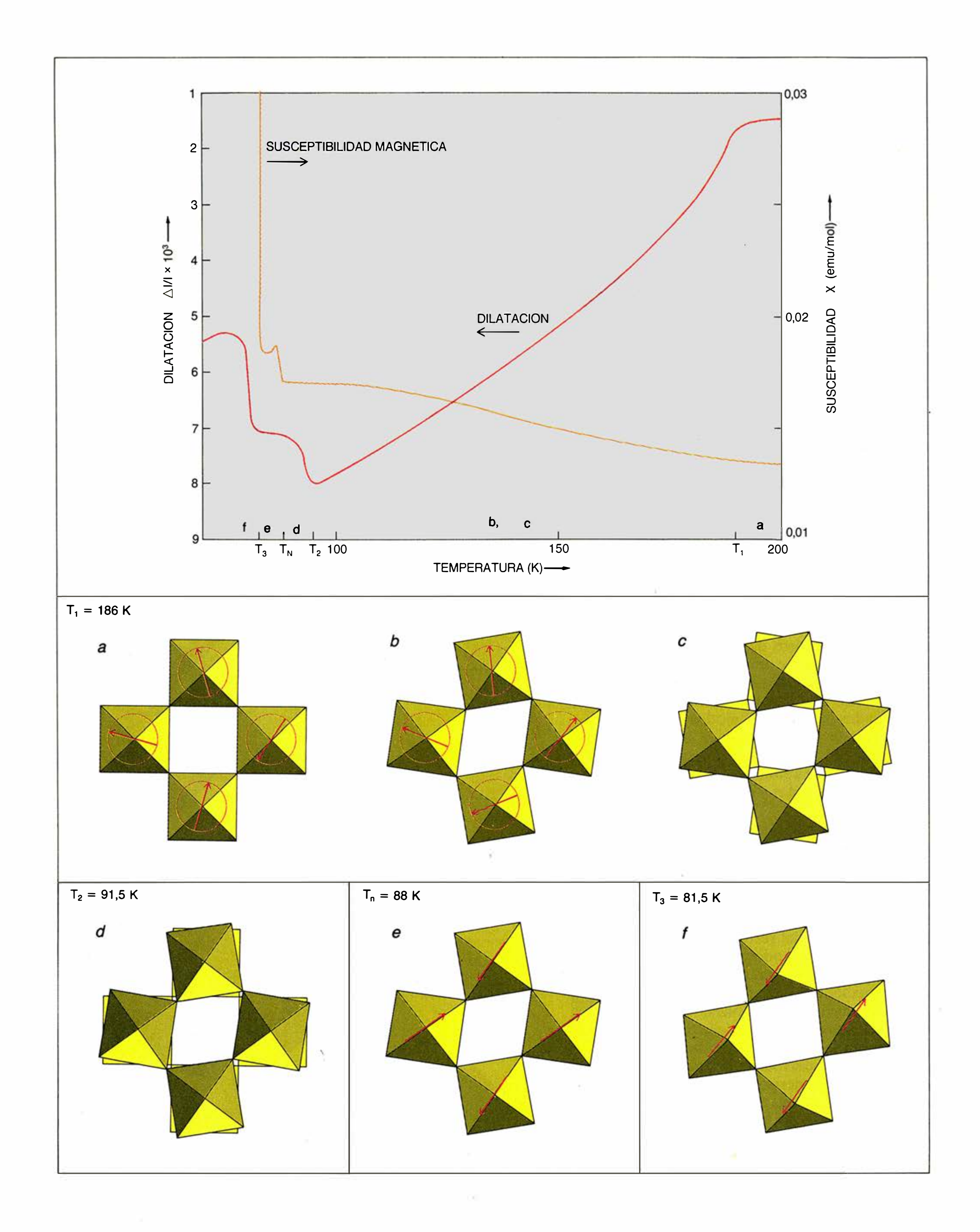
La situación más simple se da, por ejemplo, en la perovskita $KZnF_3$, que mantiene la simetría cúbica a cualquier temperatura y no sufre transición de fase alguna. Por ello, es un material ideal para contrastar sus propiedades con las del resto de la serie. La energía de los modos normales de vibración puede estudiarse directamente mediante la técnica de dispersión inelástica de neutrones. Si se bombardea con un haz de neutrones y se analiza la ganancia o pérdida de energía de los neutrones dispersados, se obtienen las energías de vibración características de los modos fonónicos. Los efectos de la absorción de energía se reflejan globalizadamente en la capacidad calorífica, cociente entre la energía térmica que se absorbe y el aumento de temperatura que provoca. Las medidas de dispersión de neutrones junto con las de capacidad calorífica han demostrado que la descripción de las propiedades físicas de esta perovskita puede hacerse mediante modos fonónicos. Además, se ha probado la práctica ausencia de




**1. ESTRUCTURA CRISTALOGRAFICA** de la perovskita cúbica genérica $ABX_3$. En la celda unidad, los cationes A y B se localizan en los vértices y en el centro del cubo, respectivamente (*esquema de la izquierda*). Seis aniones $X^-$ se enlazan fuertemente con el B formando un octaedro regular muy estable. La estructura puede entenderse como formada por octaedros regulares que, a modo de elementos de construcción, se enlazan entre sí por los vértices en las tres direcciones del espacio (*esquema de la derecha*). Los átomos A ocupan las posiciones intersticiales entre cada ocho octaedros.




**2. OSCILACIONES COLECTIVAS** (modos normales) de la red de octaedros $BX_6$. En (*a*) se ilustra el modo $R_{25}$, en que el sentido de rotación de los octaedros contiguos a lo largo del eje *c* es opuesto; en (*b*), el modo $M_3$, en que todos los octaedros a lo largo del *c* giran en el mismo sentido. Al enfriar, la frecuencia de uno de estos dos modos disminuye (se "ablanda el modo") hasta hacerse cero (cesa la oscilación) en la transición (*c*). Se ha representado en el plano $a-b$ la proyección de cuatro octaedros en fase cúbica y en la distorsionada. Si un octaedro gira alrededor del eje *c* en un sentido, los octaedros del plano que comparten un vértice lo harán, a su vez, en sentido contrario. Por último, (*d*) ofrece la estructura tetragonal resultante por condensación de un modo $R_{25}$ con el eje *c* alargado.



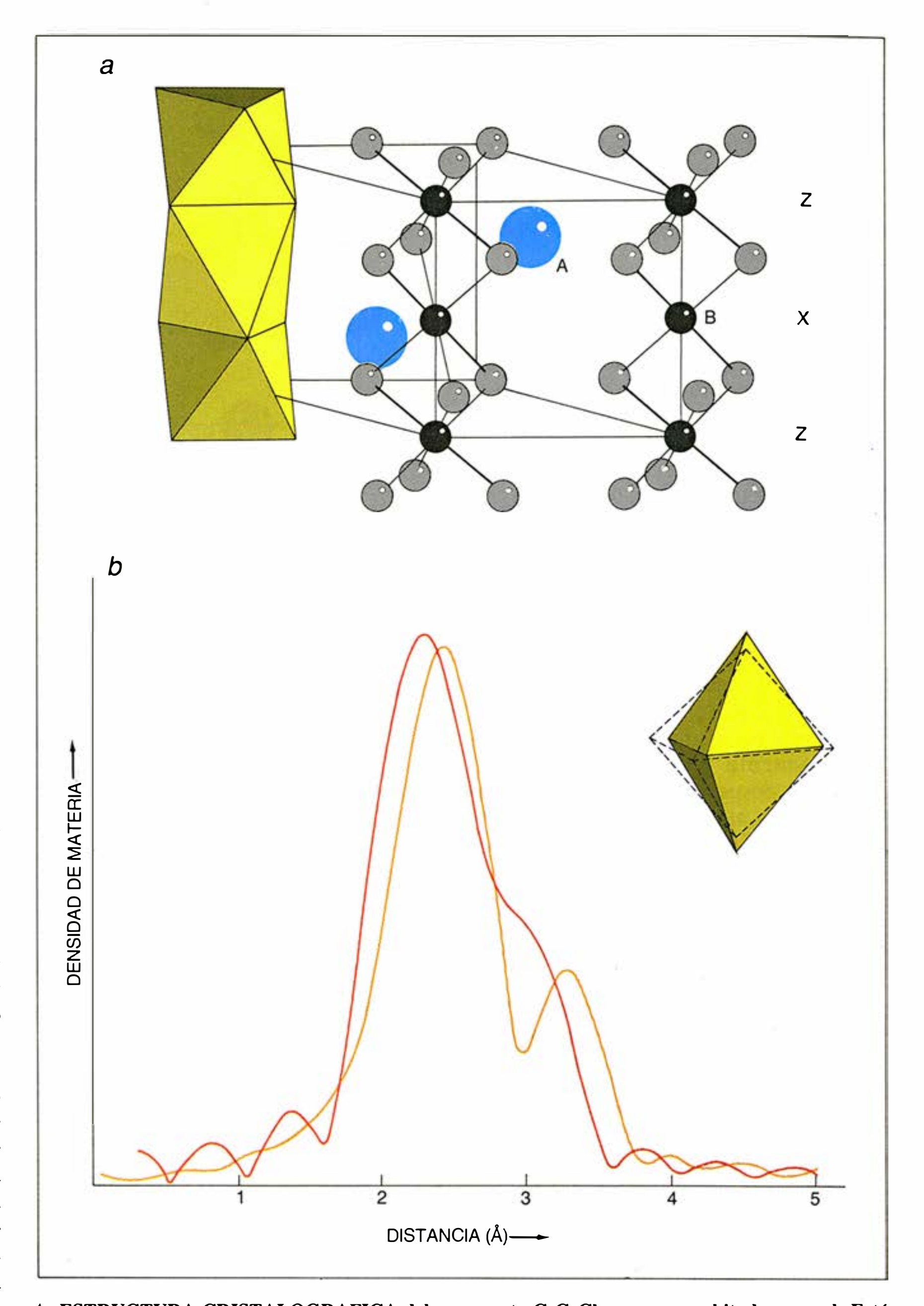

**3. TRANSICIONES DE FASE SUCESIVAS en el compuesto $KMnF_3$. En la gráfica se ha representado la variación de la dilatación términa lineal en la dirección *c* (*línea roja*) y la susceptibilidad magnética medida en la misma dirección (*línea naranja*). Las transiciones puramente estructurales se ponen de manifiesto, a $T_1$ de 186 grados K, de fase cúbica y desordenada magnéticamente (*a*) a tetragonal (*b,c*) por ablandamiento de un modo $R_{25}$ y a $T_2$ de 91,5 grados K de tetragonal a ortorrómbica (*d*) por ablandamiento del modo $M_3$. A la temperatura de Neel (Tn) de 88 grados K, se observa un pequeño pico asociado al ordenamiento anti-ferromagnético (*e*) con compensación exacta de los momentos. A $T_3$ de 81,5 grados Kelvin el acoplamiento magnetoelástico deforma la red y descompensa los momentos (*f*), dando lugar a la aparición simultánea de débil ferromagnetismo.**

anarmonicidad en el $KZnF_3$; esto es, los enlaces interatómicos actúan como si fueran muelles elásticos.

En otros casos, al bajar la temperatura, la frecuencia de un tipo de oscilación de los octaedros tiende a cero y la amplitud de dicho movimiento de vaivén disminuye hasta llegar a detenerse, quedando cada octaedro girado respecto de su posición original. Este hecho se denomina ablandamiento o condensación de un modo y permite predecir la deformación final de la estructura. La disminución y desaparición de la oscilación se debe, en general, a falta de armonicidad en los enlaces, a la existencia de más de una posición de equilibrio de los átomos o a ambas causas [*véase la figura 2c y d*]. La reducción de la simetría cúbica por ablandamiento de alguno de los modos de vibración conduce a variaciones en las posiciones de los iones; variaciones que suelen asociarse a cambios múltiples en sus propiedades.

Un ejemplo característico de la situación mencionada se observa en el $BaTiO_3$. Al disminuir la temperatura, la celda elemental pasa de simetría cúbica a tetragonal por alargamiento de una de las aristas, posteriormente se alarga una de las diagonales de las caras adquiriendo simetría ortorrómbica y, por último, se alarga una de las diagonales del cuerpo, dando lugar a simetría romboédrica. Los iones titanio ($Ti^{4+}$), partiendo de la posición que ocupan en simetría cúbica, al bajar la temperatura se desplazan en las direcciones citadas. A las temperaturas en que se producen las transiciones de fase (cambios de simetría), la permitividad eléctrica, la capacidad calorífica y la expansión térmica, entre otras magnitudes físicas, varían de forma anómala. En concreto, en la primera transición a 393 grados Kelvin aparecen dipolos eléctricos que se alinean de manera espontánea (*fase ferroeléctrica*), y la permitividad eléctrica aumenta considerablemente. Por otro lado, en las transiciones a temperaturas inferiores se producen cambios discontinuos en el valor y dirección de los momentos dipolares, aunque el material sigue siendo ferroeléctrico.

En el caso de que se produzca una gran deformación en el trascurso de una transición de fase, aparece un efecto secundario. El material se rompe o fractura en dominios cristalográficos, cuyos ejes principales en la fase deformada coinciden con los equivalentes en la fase de simetría superior. Una vez que el cristal se ha dividido en pequeños dominios o maclas, la transición de fase se produce abruptamente. Esta característica dificulta a menudo el estudio y la aplicación de estos materiales.




**4. ESTRUCTURA CRISTALOGRAFICA del compuesto $CsCrCl_3$, una perovskita hexagonal. Está formada por apilamiento de octaedros que comparten caras y forman cadenas acopladas entre sí. El eje trigonal de los octaedros $BX_6$ se alinea en la dirección del eje *c*. En cada octaedro $BX_6$ la línea gruesa representa la diagonal distorsionada por el efecto Jahn-Teller (*arriba*). El efecto Jahn-Teller produce distorsión del complejo octaédrico a lo largo de uno cualquiera de los ejes diagonales *x*, *y* o *z* (*abajo*). En la gráfica inferior se presenta la distribución radial de masa alrededor del cromo predicha para un octaedro regular (*línea naranja*) y distorsionado (*línea roja*). La medida obtenida por técnica EXAFS a alta y a baja temperatura se ajusta excelentemente con el caso distorsionado.**

## Presencia de momentos magnéticos

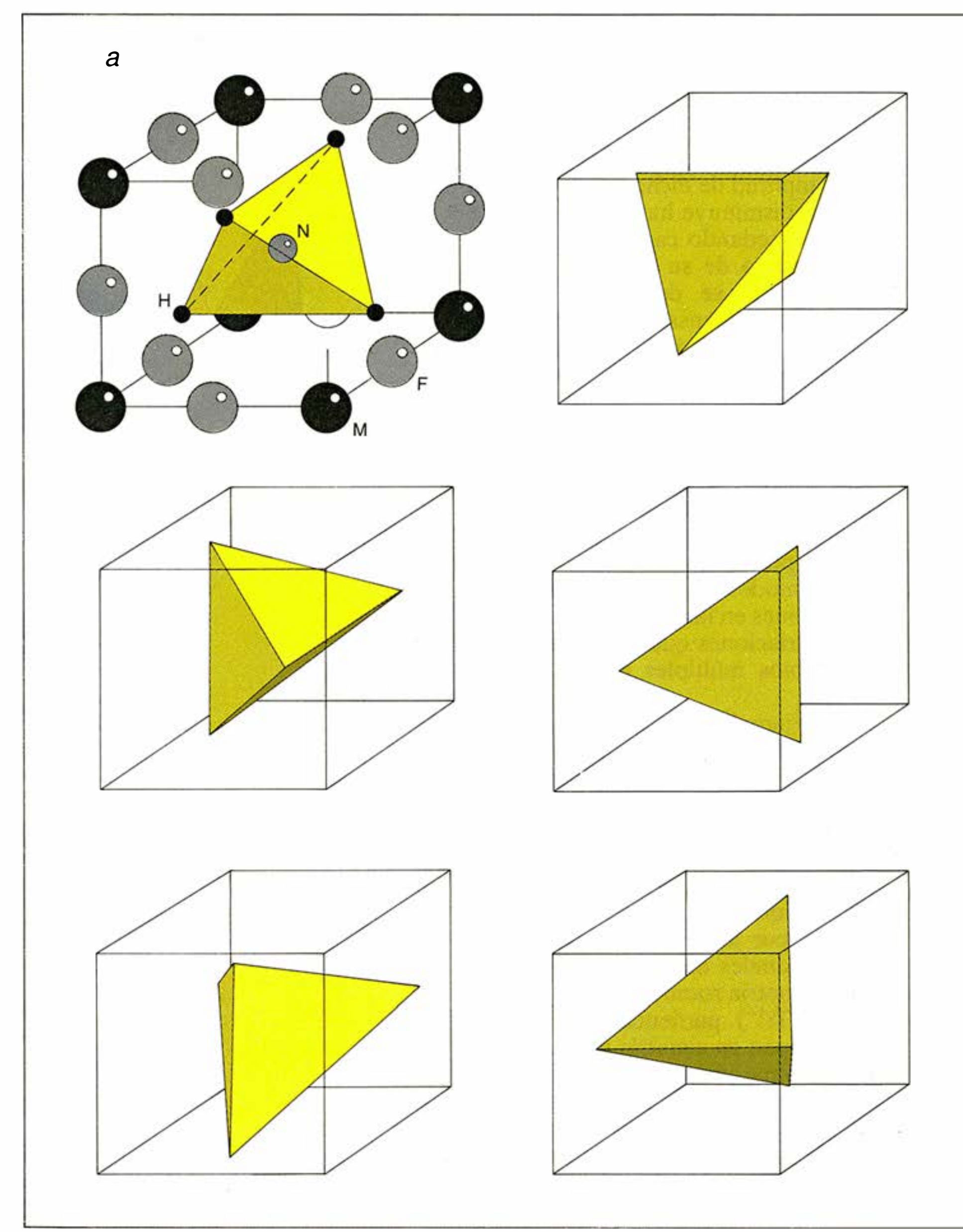
La presencia de momentos magnéticos localizados en los cationes B añade grados de libertad al sistema y posibilidades de acoplamiento con los octaedros, aumentando el número de posibles fases. El acoplamiento más débil entre la red cristalina y los momentos magnéticos se presenta en iones con estructura electrónica similar a la del ion manganeso ($Mn^{2+}$). Un prototipo de este comportamiento sería el caso del $RbMnF_3$, compuesto que tiene estructura cúbica ideal en todo el rango de temperatura. Los momentos magnéticos de los Mn próximos interaccionan por intercambio electrónico a través de puentes Mn-F-Mn (interacción de superintercambio); el efecto neto consiste en que los momentos magnéticos vecinos tienden a orientarse antiparalelamente. Por debajo de una temperatura crítica

($T_N$ = 83 grados Kelvin), llamada temperatura de Neel, la interacción induce su ordenamiento sin que haya cambio de simetría ni otras variaciones apreciables en la red cristalográfica. Hay dos subredes cúbicas, interpenetrantes, idénticas y con compensación exacta de los momentos (*antiferromagnetismo*).

En el $KMnF_3$ la sustitución del ion rubidio ($Rb^+$) por potasio ($K^+$), más pequeño, inestabiliza la red y se producen sucesivas transiciones de fase, independientemente de la interacción magnética. Partiendo de la fase cúbica de alta temperatura [*figura 3a*], la condensación del modo $R_{25}$ a 186 grados Kelvin conduce a una fase distorsionada tetragonal [*figuras 3b y 3c*]. El modo $M_3$ que sigue activo a dicha temperatura, se ablanda a 91 grados K y su distorsión se superpone a la anterior [*figura 3d*]. Ahora bien, como el sentido de la rotación del modo $M_3$ respecto del $R_{25}$ coincide en una capa y se opone en la siguiente, el ángulo total girado por los octaedros aumenta y disminuye en capas alternas, respectivamente. A temperaturas aún más bajas, 81,5 grados K, se produce una distorsión de los propios octaedros, que dejan de ser regulares. Estas transiciones se reflejan en las medidas de dilatación térmica y susceptibilidad magnética de monocristales [*figura 3*].

En la fase cúbica, los momentos están desordenados (*fase paramagnética*), las transiciones estructurales no causan efectos magnéticos detectables y el ordenamiento magnético se produce espontáneamente a 88 grados K, sin cambio estructural asociado. Sin embargo, por debajo de esta temperatura los iones $Mn^{2+}$ ocupan cuatro posiciones no equivalentes y, a su vez, cada momento magnético tiene su dirección fija dentro del octaedro. Como consecuencia de ello se forman cuatro subredes magnéticas [*figura 3e*]. A más baja temperatura se produce descompensación de los momentos debido a una ulterior distorsión, apareciendo *una fase débilmente ferromagnética* [*figura 3f*]. El acoplamiento estructural y magnético, que provoca la distorsión de los octaedros, es débil; se habla entonces de origen magnetoelástico.

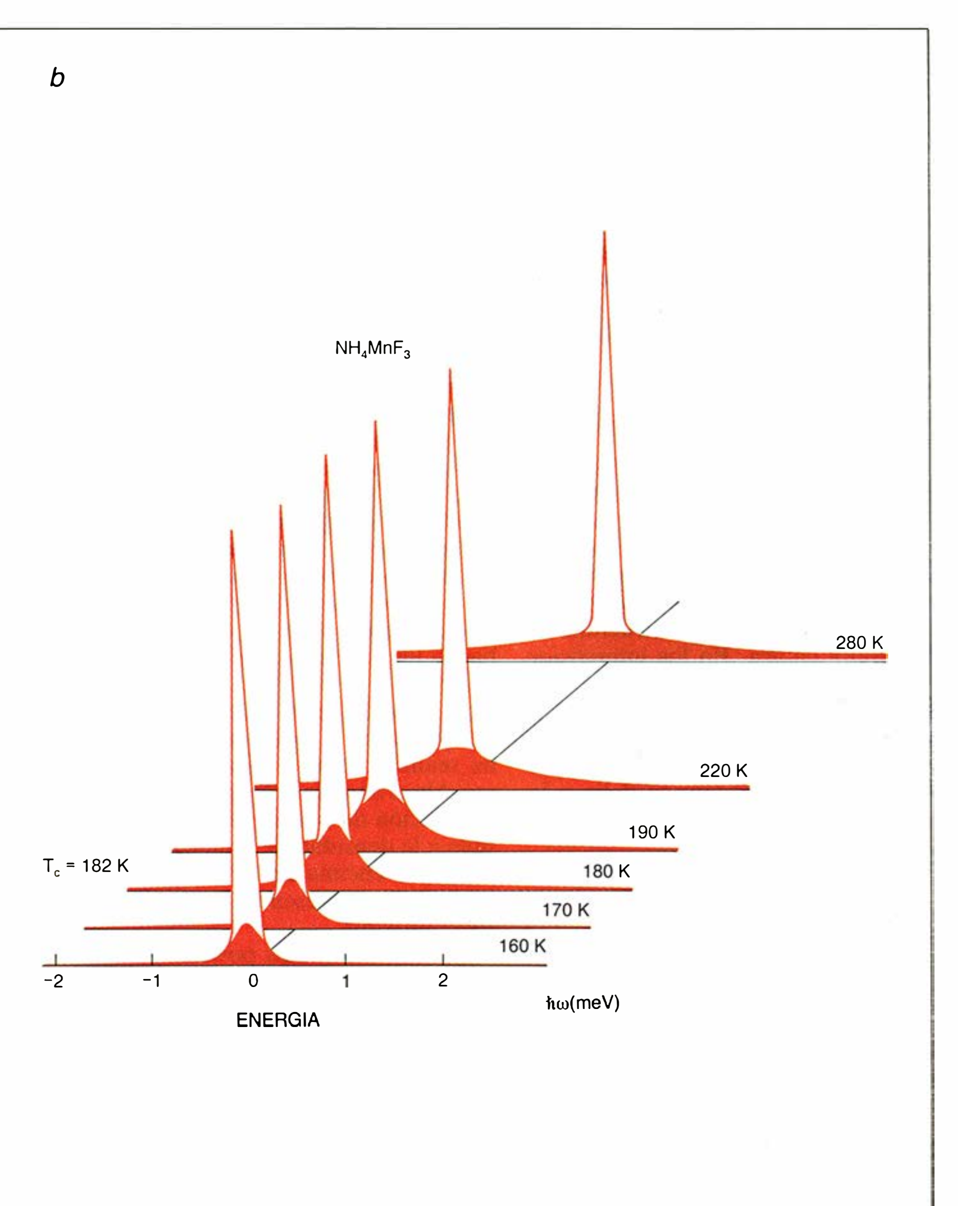
En perovskitas con iones B=$Co^{2+}$, $ACoF_3$, A=$Tl^+$, $Rb^+$, $K^+$, $NH_4^+$, el acoplamiento entre momentos magnéticos y distorsiones es más intenso. La distorsión estructural, el ordenamiento magnético y, en algunos casos, la aparición de débil ferromagnetismo son simultáneas. Este mayor acoplamiento magneto-estructural se debe a la mayor anisotropía magnética del ion cobalto ($Co^{2+}$).




**5. ORIENTACIONES DE MINIMA ENERGIA** del ion amonio, $NH_4$, en la perovskita cúbica $NH_4MF_3$ (*a*). Los hidrógenos apuntan hacia los iones $F^-$ debido al enlace por puente de hidrógeno N-H...F. Cuando se excita el ion $NH_4^+$ salta de una a otra orientación estocásticamente. En *b*, la

Un mayor acoplo electrónico-estructural se presenta en perovskitas con iones B=$Cu^{2+}$ y $Cr^{2+}$; el cobre y el cromo poseen, respectivamente, un electrón o un hueco en su última capa electrónica. El entorno octaédrico disminuye su energía alargando una de sus diagonales, fenómeno que constituye el efecto Jahn-Teller estático, que debe su nombre a H. A. Jahn y E. Teller. Este es el caso de las perovskitas $KCuF_3$ y $KCrF_3$, cuya estructura es tetragonal porque, de partida, los octaedros individuales $BF_6$ están deformados a cualquier temperatura. Los octaedros, al compartir vértices, se conectan con las diagonales alargadas perpendiculares entre sí, dando simetría tetragonal. En estos dos compuestos, el efecto Jahn-Teller es tan intenso que la deformación se mantiene a cualquier temperatura y, por tanto, no hay transiciones de fase estructurales.

Por el contrario, otras perovskitas de cobre y cromo, como los compuestos $CsCuCl_3$, $CsCrCl_3$ y $RbCrCl_3$, presentan deformaciones dinámicas que dan lugar a transiciones de fase, y hablamos entonces de efecto Jahn-Teller dinámico. En estas perovskitas los octaedros $BCl_6$ se apilan compartiendo caras [*véase la figura 4a*]; la simetría es hexagonal. La red cristalográfica se considera formada por cadenas paralelas de octaedros separadas por iones Cs o Rb. El efecto Jahn-Teller activa el alargamiento de una




**componente casi-elástica de la dispersión incoherente de neutrones (*zona roja de los espectros*) se ensancha al aumentar la temperatura por encima de la de transición ($T_c$ = 182 K) y pone de manifiesto que los grupos $NH_4^+$ pasan de una a otra orientación cada vez con mayor frecuencia.**

diagonal por octaedro, pero a altas temperaturas es posible que resulte alargada una cualquiera de las tres diagonales equivalentes y perpendiculares X, Y, Z. El efecto resultante es una deformación dinámica al azar, como lo confirman las medidas de espectroscopía visible, infrarroja y EXAFS ("Extended X-Ray Absortion Fine Structure) [*figura 5b*]. Ahora bien, la dirección de deformación de cada octaedro cambia rápidamente y la difracción de rayos-X, que promedia en el tiempo las posiciones atómicas, detecta octaedros regulares.

Al disminuir la temperatura, se dificulta hasta acabar por cesar; cesa el cambio entre las direcciones de deformación, hasta que se detiene el movimiento. Todos y cada uno de los octaedros quedan distorsionados en una sola de sus diagonales, fenómeno éste conocido por *transición Jahn-Teller cooperativa*. La estructura de baja temperatura corresponderá a la organización espacial de octaedros deformados que tenga menor energía elástica. El $CsCrCl_3$ a 171 grados K se deforma y pasa a simetría monoclínica, en la que las diagonales deformadas de los octaedros a lo largo de las cadenas crean una secuencia ZXZYZX. En el compuesto $CsCuCl_3$, la transición a 423 grados K conduce a una secuencia ZXYZXY, mientras que en el $RbCrCl_3$ hay una estructura intermedia Z(XY)Z. El paréntesis significa que, fijada la distorsión de un octaedro a lo largo de la diagonal Z, los adyacentes se deforman dinámicamente entre X e Y dando una estructura parcialmente desordenada. Este desorden estabiliza la fase intermedia, ya que cualquiera de las secuencias tiene la misma energía elástica.

## Transiciones de reorientación molecular

Los movimientos de los octaedros $BX_6$ y de los iones A también están acoplados y se manifiestan con mayor intensidad cuando estos últimos tienen grados de libertad adicionales. Un caso significativo se produce en perovskitas que contienen el ion amonio ($NH_4^+$), verbigracia, la $NH_4MF_3$. Los átomos H del $NH_4^+$ forman un tetraedro regular que gira como un todo alrededor de su centro de masas, o bien elige orientaciones preferentes cuando se acopla con la red cristalina. En particular, el enlace H-F es muy intenso, por lo que los H apuntan hacia los $F^-$. A causa de la simetría cúbica, existen seis orientaciones posibles y geométricamente distinguibles con la misma energía de enlace. En virtud de lo cual, el ion $NH_4^+$ pasa de una a otra si posee energía suficiente para saltar la barrera energética existente entre ellas [*véase la figura 5a*].

Con difracción elástica de neutrones se ha determinado que los átomos de hidrógeno de cada $NH_4^+$ residen en 24 posiciones situadas simétricamente dos a dos sobre las líneas N-F y que simultáneamente sólo se ocupan de cuatro en cuatro. Sin embargo, todas tienen la misma probabilidad de ser visitadas por los hidrógenos, y así lo hacen de forma dinámica mediante reorientaciones del $NH_4^+$ entre las seis posiciones posibles.

Mediante dispersión cuasi-elástica de neutrones se puede medir la ganancia o pérdida de energía cuando éstos colisionan con los átomos H del $NH_4^+$ en rotación. El gran número de colisiones producidas da una estadística indicativa de los movimientos más frecuentes del $NH_4^+$, que son giros alrededor de los ejes cuaternarios del cubo. Además, la disminución exponencial de la frecuencia del salto con la temperatura, que se manifiesta en la figura *5b* como un ensanchamiento del espectro cuasi-elástico, permite obtener la energía de la barrera que debe superar el ion amonio para reorientarse.

La observación directa de la dependencia de las "alturas" de las barreras de energía con la distancia puede realizarse aplicando altas presiones hidrostáticas, que reducirán progresivamente el volumen [véase "Celdilla

de yunques de diamante para producir altas presiones", por A. Jayaraman; INVESTIGACIÓN Y CIENCIA, junio de 1984]. Las medidas de espectroscopía Raman muestran un incremento importante en la energía de las vibraciones de los H en la dirección del enlace N-H al aumentar la presión, en contra de lo esperado si el enlace fuera exclusivamente electrostático. Esta experiencia indica, por tanto, que las fuerzas interatómicas deben tener una fuerte componente repulsiva.

Una característica de la serie $NH_4MF_3$ es la presencia de una anomalía aguda en la capacidad calorífica a la temperatura de la transición de reorientación, cuya entropía es muy superior a las observadas en otras perovskitas que no contienen amonio. Curiosamente, el contenido entrópico es próximo al logaritmo neperiano de 3. La variación de entropía se relaciona con el logaritmo del cociente del número de orientaciones posibles antes y después de la transición. Por tanto, la presencia de seis orientaciones en la fase de alta temperatura implicaría que en la de baja deberán existir dos orientaciones, lo que se ha confirmado en medidas del espectro Raman de las vibraciones características del $NH_4^+$.

De todo lo anterior resulta evidente la existencia de un acoplamiento fuerte entre la subred de $NH_4^+$ y los octaedros $MF_6^-$ tal que, cuando éstos se deforman, arrastran a los iones amonio y provocan una reducción en sus grados de libertad previos. Ese fenómeno se denomina *transición de reorientación*.

## Perovskitas conductoras

Las perovskitas metálicas $Mn_3MN$ derivan estructuralmente del $Mn_4N$. El nitrógeno se enlaza fuertemente con los seis Mn más próximos y se constituye un octaedro regular muy estable. Los octaedros comparten vértices y forman así la estructura de perovskita cúbica. En los intersticios entre octaedros se localizan átomos de Mn, más débilmente enlazados y de fácil sustitución por distintos metales de transición. El fuerte enlace N-Mn en los octaedros se debe a la hibridación de orbitales electrónicos localizados *p* del N y *d* del Mn, mientras que el carácter metálico se debe a la banda de estados electrónicos *s* del Mn y *s, p* de los metales intersticiales M.

Los modos de vibración colectiva de estos octaedros son los mismos que en las perovskitas aislantes, pero el problema es más complejo porque hay bandas de electrones localizados (bandas *d* puras e híbridas *p-d*) y deslocalizados (bandas *s* y *p*) que se superponen en energía. Los electrones ocupan todos los estados de las bandas hasta llegar a la energía de Fermi, que es la máxima que puede tener un electrón a la temperatura del cero absoluto. La energía de Fermi es sensible a la forma de las bandas y al número de electrones que aportan los metales de transición. En particular, este número puede ser muy alto si la banda localizada proveniente de la hibridación *p-d* tiene una energía próxima a la de Fermi. Una comprobación indirecta de esta proposición se ha realizado con técnica XANES ("X-ray Absorption Near Edge Structure") y radiación de sincrotrón, que es sensible a la densidad de estados electrónicos por encima del nivel de Fermi, detectando la presencia de una banda vacía próxima.

La simetría cúbica posibilita la triple degeneración espacial: tres direcciones equivalentes y tres estados electrónicos con energía idéntica; además, cada electrón dispone de dos estados espinoriales distintos correspondientes a los dos valores posibles del número cuántico de espín ($S_z = \pm 1/2$), y la degeneración total es seis. Cualquier interacción que rompa la degeneración, esto es, que minimice la energía de algunos estados, provoca una transición. Por ejemplo, a 1500 grados K el $Mn_4N$ pasa de simetría hexagonal a cúbica y a 756 grados K se ordena magnéticamente. En esta última transición se rompe la degeneración espinorial y se produce un ordenamiento *ferrimagnético*, esto es, los momentos magnéticos que son distintos en los dos tipos de Mn se alinean antiparalalelamente.

En las perovskitas $Mn_3MN$, donde M representa átomos de galio (Ga), cinc (Zn), cobre (Cu) o níquel (Ni), que no poseen momento magnético, el ordenamiento debería ser más simple, porque sólo intervienen los momentos de los Mn. Sin embargo, aparecen efectos debidos a *frustración* en el ordenamiento de los momentos. Así, a 278 grados K en el $Mn_3GaN$ y a 191 grados K en el $Mn_3ZnN$ hay un ordenamiento antiferromagnético triangular, con aumento de volumen y sin cambios de simetría.

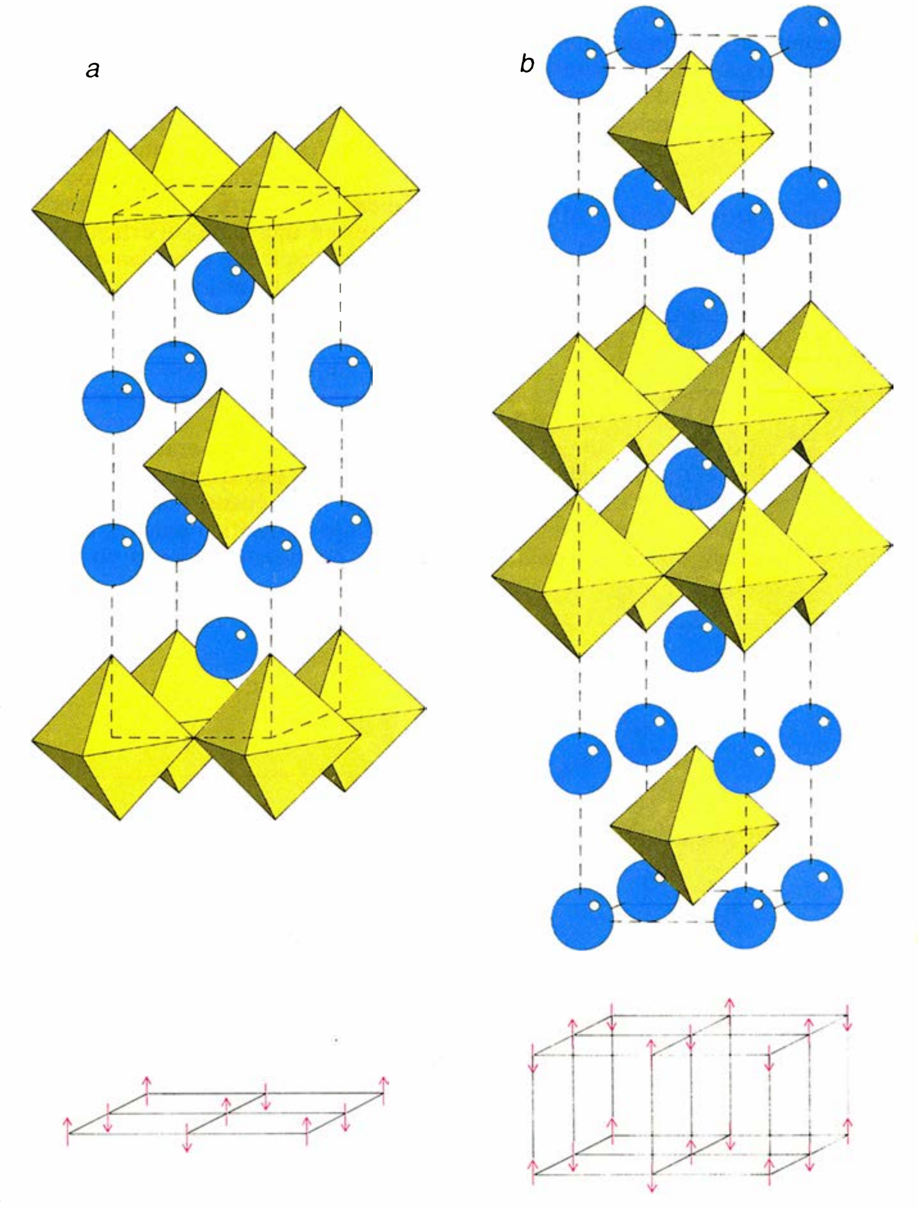


**6. PEROVSKITA METALICA $Mn_3MC$. Arriba se ilustra la capacidad calorífica; muestra anomalías en la transición de fase paraferromagnética a $T_N = 243$ K y ferro-antiferromagnética a $T_{F\text{-}AF} = 171{,}8$ grados K. Abajo se dibujan la celda elemental y estructuras magnéticas en las dos fases ordenadas.**

En el $Mn_3ZnN$, y a 127 grados K, hay una segunda transición; consiste ésta en una contracción de volumen, cambio de estructura magnética y reducción en el valor de los momentos magnéticos, que se interpreta como un incremento en la deslocalización de los electrones de valencia. La diferencia esencial entre ambos compuestos radica en que el Ga aporta un electrón más que el Zn, lo que conduce a la disminución del volumen de la perovskita cúbica en la fase paramagnética.

En la sustitución de nitrógeno por carbono, las diferencias esenciales se deben a que este último aporta un electrón menos que el primero y a que la energía del enlace híbrido es menor. Considerando carburos y nitruros isoelectrónicos, tales como $Mn_3GaC$ y $Mn_3ZnN$ (el electrón en exceso del Ga es compensado por el defecto del C), su comportamiento debiera ser semejante. Sin embargo, la energía de los electrones *p* en los carburos es inferior a la del orbital *d* del Mn que se encuentra en resonancia en el enlace híbrido, dando como consecuencia estructuras magnéticas distintas. Hemos ilustrado en la figura 6 las anomalías en la capacidad calorífica producidas a una temperatura de Neel de 243 grados K por el ordenamiento ferromagnético de los momentos de los Mn en dirección de la diagonal principal del cubo, y una transición de reorientación ferro-antiferromagnética a 172 grados K.

El llenado electrónico de la banda de valencia puede hacerse de forma continua, sustituyendo átomos de Mn intersticial por cinc o galio en las disoluciones sólidas $Mn_3(Mn_xM_{1-x})N$. En su diagrama de fase estructural y magnético aparece una fase magnética parcialmente desordenada denominada vidrio espinorial. En ella los momentos magnéticos presentan viscosidad a la rotación cuando se los somete a campos magnéticos.




**7. ESTRUCTURAS DERIVABLES de la perovskita cúbica. A la izquierda, en el $Ca_2MnO_4$ los octaedros $MnO_6$ comparten cuatro vértices formando capas de octaedros separados por átomos de Ca. A la derecha, en el $Ca_3Mn_2O_7$, los octaedros comparten cinco vértices formando capas dobles de octaedros. El ordenamiento de los momentos aparece en la fase ordenada anti-ferromagnética.**

## Oxidos con estructura de perovskita

Aunque hemos analizado las transiciones de fase de las perovskitas aislantes y conductoras por separado, en los óxidos ($ABO_3$) no puede aplicarse esta división, ya que hay perovskitas como $SrVO_3$ y $LaNiO_3$ que tienen conducción metálica; otras, por ejemplo, $LaFeO_3$ y $LaCrO_3$, son semiconductoras, y otras, $SrTiO_3$ y $BaTiO_3$, se comportan como buenos aislantes.

Además, determinadas perovskitas se convierten en superconductoras (verbigracia $BaPb_{1-x}Bi_xO$)[3], aunque a temperaturas por debajo de los 13 grados K. Esta variedad, unida a las distorsiones estructurales, da una gama amplia de propiedades de interés y de transiciones de fase. Describimos tan sólo algunos de los óxidos menos conocidos en cuanto a sus propiedades y estructura.

## Estructuras derivadas de la perovskita

Todo el análisis estructural anterior se ha basado en la consideración de los octaedros como poliedros básicos conectados por sus seis vértices. Partiendo de ellos pueden aparecer estructuras similares en las que los octaedros poseen en común un número menor de vértices. Por ejemplo, en el $Ca_2MnO_4$ los octaedros $MnO_6$ sólo comparten cuatro vértices, formando capas separadas por otras de iones $Ca^{2+}$, siendo la simetría tetragonal [*véase la figura 7a*]. Con esta geometría ha sido posible estudiar la dependencia de las propiedades magnéticas de un sistema cuasibidimensional gracias a la interacción, muy débil, entre capas.

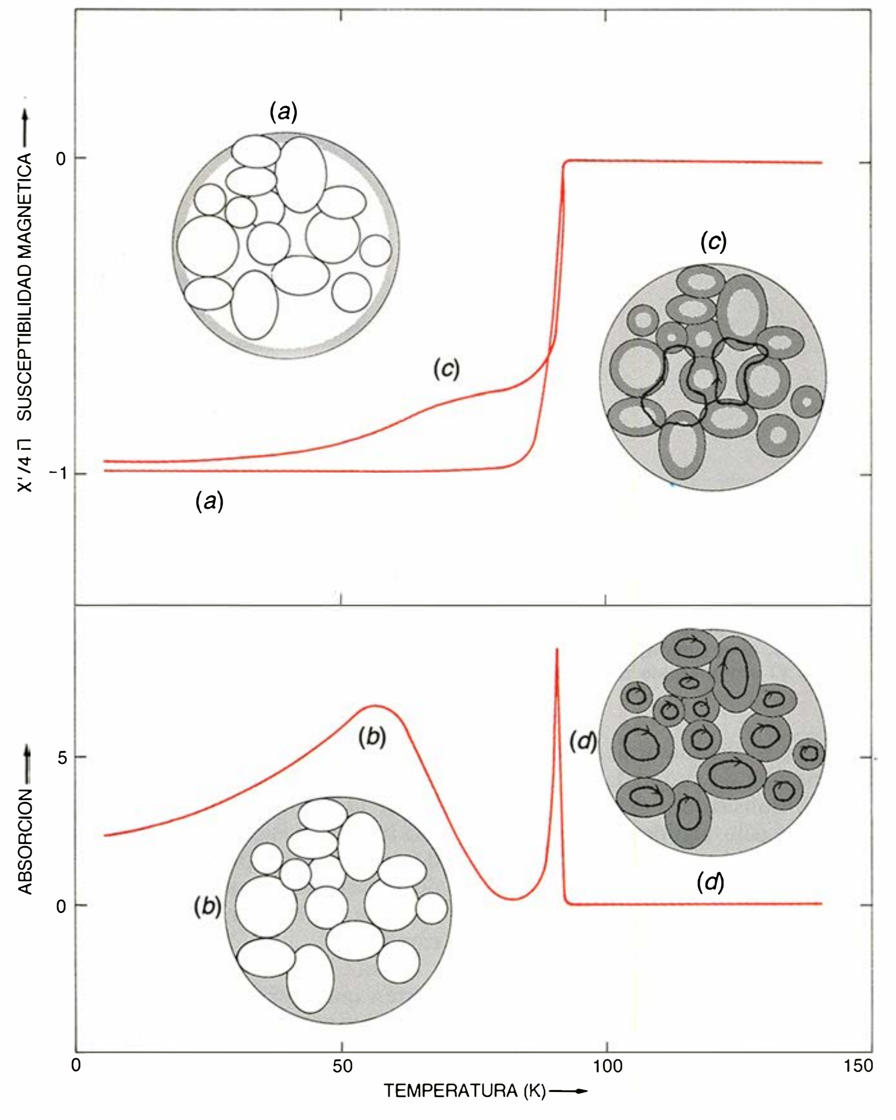
En el $Ca_3Mn_2O_7$ los octaedros comparten cinco vértices y forman capas dobles separadas por otras de $Ca^{2+}$, como en el caso anterior. En el $Ca_4Mn_3O_{10}$ dos tercios de los octaedros comparten cinco vértices y un tercio seis, formándose así capas triples. Caben otras muchas combinaciones, pero todas tienen el hecho estructural común de contener la celda elemental de las perovskitas cúbicas, lo que permite su análisis siguiendo

las líneas vistas hasta ahora. Mediante esta secuencia de compuestos se ha podido observar el "cruzamiento" de dimensionalidad desde el caso bidimensional al tridimensional [*véase la figura 7b*].

## Superconductores de alta temperatura crítica

Los superconductores de alta temperatura crítica provienen de las perovskitas $ABO_3$. La celda elemental ortorrómbica de la serie $RBa_2Cu_3O_{7-\delta}$ (donde R sustituye al itrio o una tierra rara) está formada por una celda cúbica unidad de $RCuO_3$, situada entre dos de $BaCuO_3$; en dicha celda unitaria hay un defecto de más de dos átomos de oxígeno por fórmula unidad. Al igual que en las perovskitas cúbicas, estos nuevos superconductores sufren transiciones de fase estructurales y aparecen dominios cristalográficos en que el par de ejes *a* y *b* están orientados en distintas direcciones, esto es, se forman maclas.

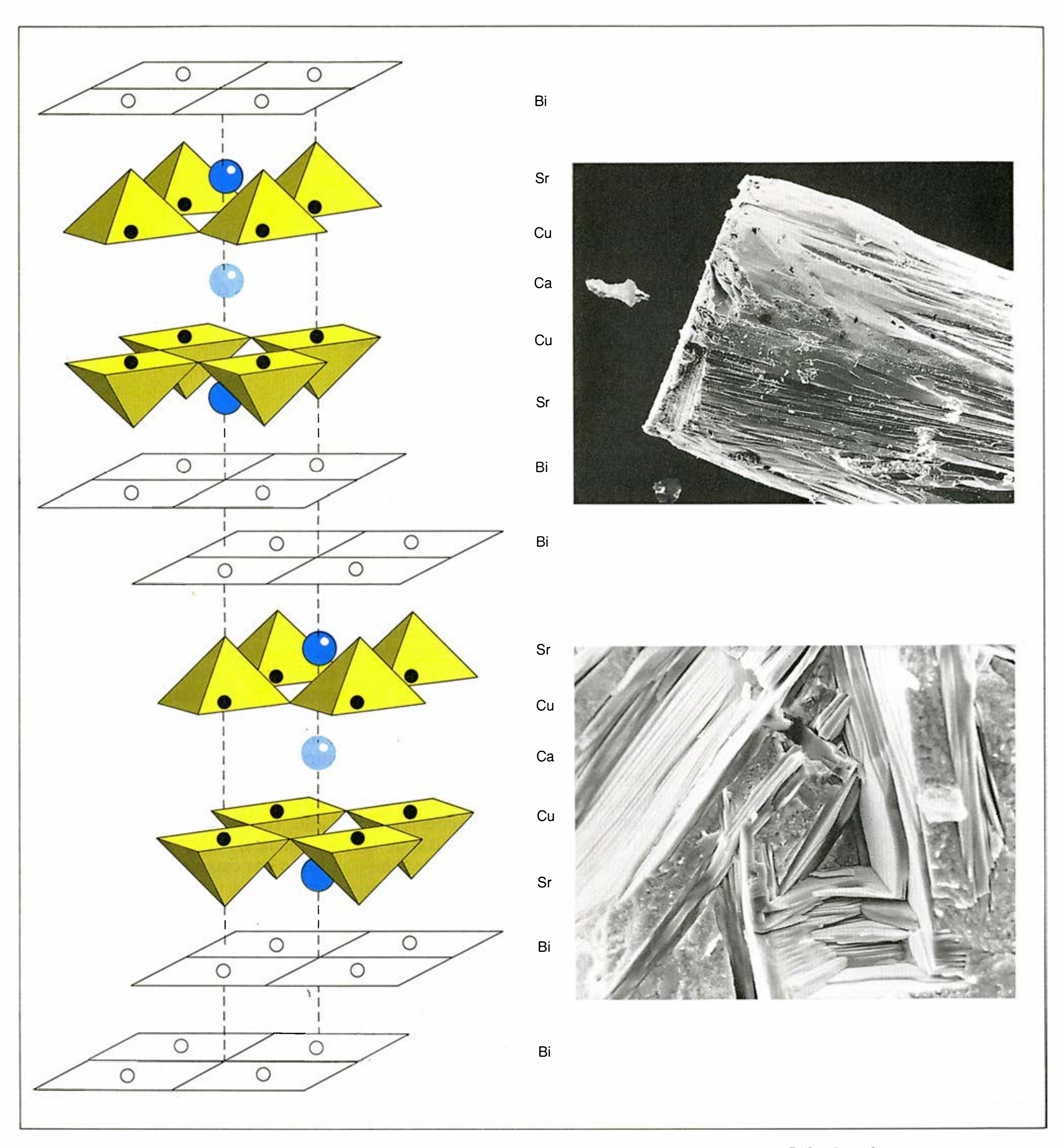
Al aplicar un campo magnético a un superconductor se inducen supercorrientes en una capa superficial de pequeño espesor que expulsan el flujo magnético del interior del material. Este efecto recibe el nombre de su descubridor, W. Meissner. La expulsión de flujo es completa si dicho campo es inferior a $H_{c1}$ (campo crítico inferior). Ahora bien, los superconductores de alta temperatura crítica son de tipo II; ello quiere decir que, a campos magnéticos mayores que $H_{c1}$, el flujo magnético penetra en el material, existiendo regiones en forma de tubos que están en fase normal, rodeadas de torbellinos de corrientes superconductoras. Los torbellinos pueden moverse transversalmente por la acción de campos magnéticos y de corrientes eléctricas. Las tensiones locales, las fronteras de grano y otros defectos del cristal tienden a fijar los tubos de flujo y determinan los valores de las corrientes y campos críticos, $H_{c2}$, para los cuales el material pasa al estado normal. El diamagnetismo asociado es intenso, fácilmente detectable con técnicas de medidas de susceptibilidad magnética alterna [*véase la figura 8*], permitiendo la determinación de temperaturas y corrientes críticas. Con frecuencia estos materiales son cerámicos, formados por granos cristalinos en contacto (sinterizados); a los efectos anteriores hay que añadir los que originan las uniones entre granos. En este caso, las corrientes eléctricas que pueden soportar son menores, con lo que quedan limitadas sus posibles aplicaciones.




**8. SUSCEPTIBILIDAD MAGNETICA para excitación alterna de una muestra policristalina del $YBa_2Cu_3O_{7-d}$. A muy baja temperatura (*a*), las supercorrientes circulan por la capa externa de la muestra (*zona sombreada*) sin penetrar en los granos, apantallando el interior de la muestra. Al aumentar la temperatura (*b*), el campo penetra hasta el centro de la muestra, pero los granos siguen apantallados. Las supercorrientes circulan por una capa superficial de los granos y cruzan de uno a otro a través de sus uniones (*c*). El campo penetra en los granos y se producen corrientes intragranulares (*líneas gruesas en el dibujo*). Finalmente, el material se hace normal a 91 grados Kelvin.**

La búsqueda de materiales con temperatura crítica próxima a temperatura ambiente y que permitan producir hilos superconductores susceptibles de ser utilizados en aplicaciones industriales han conducido a nuevos materiales; tal es el caso de $Bi_2Sr_2CaCu_2O_{8+\delta}$ [*véase la figura 9*]. Este material tiene una temperatura crítica de 110 grados K, razonablemente alta; su interés proviene del éxito en la síntesis de una fibra mediante las técnicas empleadas para la fabricación de fibra óptica. Se abre así un prometedor camino en la producción de materiales con textura favorable para el soporte de altas corrientes.

En el marco de la teoría BCS (J. Baarden, L. N. Cooper y J.R. Schrieffer) se sabe que la superconductividad convencional resulta de la formación de pares de electrones (pares de Cooper) mediante el acoplamiento de los electrones con las vibraciones de la red. De esta forma las cargas pueden moverse en el material sin impedimento y, por ello, el material superconductor carece de resistencia eléctrica. La mayor parte de la fenomenología de los nuevos superconductores conocida hasta la fecha puede explicarse con la teoría BCS. Se han invocado algunos de los mecanismos físicos mencionados en el artículo –acoplamiento Jahn-Teller, rotura de la degeneración espinorial de las bandas– para explicar la formación de pares de Cooper, pero hoy en día todavía se desconoce el mecanismo que actúa en este caso y qué relación existe con la estructura básica de las perovskitas.




**9. SUPERCONDUCTOR de alta temperatura crítica $Bi_2Sr_2CaCu_2O_{8+d}$, cifrada en 86 grados Kelvin. A la izquierda se esquematiza la estructura formada por capas de pirámides $CuO_5$ que comparten cuatro vértices, intercaladas por capas cuadradas de $BiO_4$, Sr y Ca. A la derecha, se da la fibra obtenida por fusión con láser (*arriba*) y se ofrece la sección transversal mostrando la microestructura de laminillas orientadas en abanico (*abajo*).**

## A modo de conclusión

Considerar en primer lugar, de forma simplificada, los efectos de las interacciones más intensas en los materiales y, posteriormente, "afinar" el modelo atendiendo a las interacciones secundarias es un método científico común que permite abordar problemas muy complejos de áreas aparentemente dispares e inducir, a partir de una extensa fenomenología, los mecanismos que lo producen.

A lo largo de este artículo se ha ilustrado este método considerando que la contribución de los octaedros $BX_6$ a la energía de formación de los compuestos es dominante. Con esta hipótesis, las vibraciones de los propios octaedros, la interacción con los momentos magnéticos, la acción de grupos moleculares con grados de libertad interna o la presencia de electrones de conducción contribuirán a la energía dando una pequeña perturbación. La pérdida de simetría y el cambio de propiedades son consecuencia de la minimización de la energía total del material.

BIBLIOGRAFIA COMPLEMENTARIA


PRINCIPLES AND APPLICATIONS OF FERROELECTRICS AND RELATED MATERIALS. M. E. Lines y A. M. Glass. Clarendon Press, 1977.

PEROVSKITAS. Robert M. Hazen en *Investigación y Ciencia*, n.° 143, págs. 46-54; agosto de 1988.

# La retroacción positiva en la economía

*Una nueva teoría económica aclara los mecanismos por los que pequeños acontecimientos aleatorios que se produzcan en los albores de una industria o tecnología deciden hacia dónde se inclinará la balanza competitiva*

W. Brian Arthur

La teoría económica convencional se basa en el supuesto de los rendimientos decrecientes. Las acciones económicas engendran una retroacción negativa que conduce a un equilibrio predecible de los precios y al reparto del mercado. Tal retroacción tiende a estabilizar la economía porque todo cambio importante se verá compensado por las reacciones mismas que genera. El alza de los precios del petróleo en el decenio de 1970 fomentó la conservación de la energía e impulsó la prospección, precipitando una caída predecible de los precios poco después de 1980. Según la teoría convencional, el equilibrio conduce al "mejor" resultado posible en cada caso, esto es, a la máxima eficiencia en el uso y asignación de los recursos.

Este cuadro tan agradable no se compadece a menudo con la realidad. En muchas partes de la economía no parece que actúen las fuerzas estabilizadoras. En su lugar, una retroacción positiva amplía los efectos de pequeños cambios económicos; los modelos económicos que describen tales efectos difieren considerablemente de los convencionales. Los rendimientos decrecientes implican un solo punto de equilibrio en la economía, cuando una retroacción positiva, es decir, unos rendimientos crecientes, permiten muchos puntos posibles de equilibrio. Nada garantiza que el resultado económico concreto seleccionado entre las muchas alternativas sea el "mejor". Asimismo, una vez que acontecimientos aleatorios hagan que la economía discurra por una determinada trayectoria, esa opción puede resultar obligada, con independencia de las ventajas de las vías alternativas. Si un producto o nación en un mercado competitivo se pone a la cabeza por "suerte", tenderá a seguir en esa posición e incluso a incrementar su ventaja. Deja de estar así garantizado un reparto del mercado predecible.

En los últimos años, teóricos de la economía de la Universidad de Stanford, del Instituto de Santa Fe de Nuevo México y de otros lugares hemos estado trabajando en una visión de la economía basada en la retroacción positiva. La economía de los rendimientos crecientes tiene antecedentes que se remontan 70 años o más, pero su aplicación a la economía en su conjunto constituye, en gran parte, una novedad. La teoría tiene mucho parecido con la moderna física no lineal (en oposición a los modelos físicos anteriores al siglo XX en que se basa la economía convencional), requiere técnicas matemáticas nuevas y sugestivas y parece ser la teoría apropiada para comprender la economía moderna de la alta tecnología.

W. BRIAN ARTHUR desempeña la cátedra Morrison de estudios de población y economía en la Universidad de Stanford. Se doctoró, en 1973, por la de California en Berkeley. Hasta hace poco, Arthur gozaba de una venia para trabajar en el Instituto de Santa Fe, un centro de investigación dedicado a estudiar sistemas complejos. Allí dirigió un equipo de economistas, físicos, biólogos y otros especialistas que investigó el comportamiento de la economía como un complejo sistema en evolución.

La historia de los magnetoscopios o grabadoras de cintas magnéticas ("videocassettes") brinda un sencillo ejemplo de retroacción positiva. El mercado de ese producto comenzó con dos formatos competitivos que se vendían más o menos al mismo precio: VHS y Beta. Ambos podían obtener rendimientos crecientes cuando su parte de mercado aumentaba: un mayor número de magnetoscopios VHS incitaría a las tiendas de vídeo a disponer de más cintas grabadas en formato VHS, incrementando así el valor de poseer un aparato VHS y haciendo que más gente adquiriera uno. (Lo mismo, claro está, se aplicaría al formato Beta.) De esta suerte, un pequeño incremento de la parte del mercado mejoraría la posición competitiva de un sistema y le ayudaría a acrecentar su ventaja.

Un mercado de esa índole resulta inicialmente inestable. Ambos sistemas se introdujeron, más o menos, al mismo tiempo y comenzaron, por tanto, con partes del mercado bastante iguales; esas partes fluctuaron pronto por causa de circunstancias externas, "suerte" y maniobras empresariales. Llegó un momento en que los rendimientos crecientes de las mejoras iniciales inclinaron la competencia a favor de VHS, que sacó los cuerpos de ventaja como para hacerse prácticamente con todo el mercado de magnetoscopios. Con todo, cuando comenzó la competencia, hubiera sido imposible decir qué sistema triunfaría, esto es, cuál de las dos posibles equilibrios sería el elegido. Asimismo, si fuera verdad lo que se aduce acerca de que Beta era técnicamente superior, entonces la elección hecha por el mercado no condujo al mejor resultado económico.

La teoría económica convencional adopta una visión diferente de la competencia entre dos tecnologías o productos que desempeñan la misma función. Un ejemplo es la competencia entre agua y carbón para generar electricidad. Conforme las centrales hidroeléctricas se hace con una parte mayor del mercado, los ingenieros tendrán que recurrir a emplazamientos más costosos para la construcción

de presas, incrementando así las posibilidades de que una central térmica resulte menos costosa. Cuando estas últimas amplían su parte del mercado, subirá el precio del carbón (o se impondrán unos costosos controles de la contaminación) y así volverá a inclinarse la balanza hacia la hidroelectricidad. Las dos tecnologías terminan repartiéndose el mercado en una proporción predecible, que aprovecha al máximo las posibilidades de cada una, en contraste con lo que ocurrió con los dos sistemas de magnetoscopios.

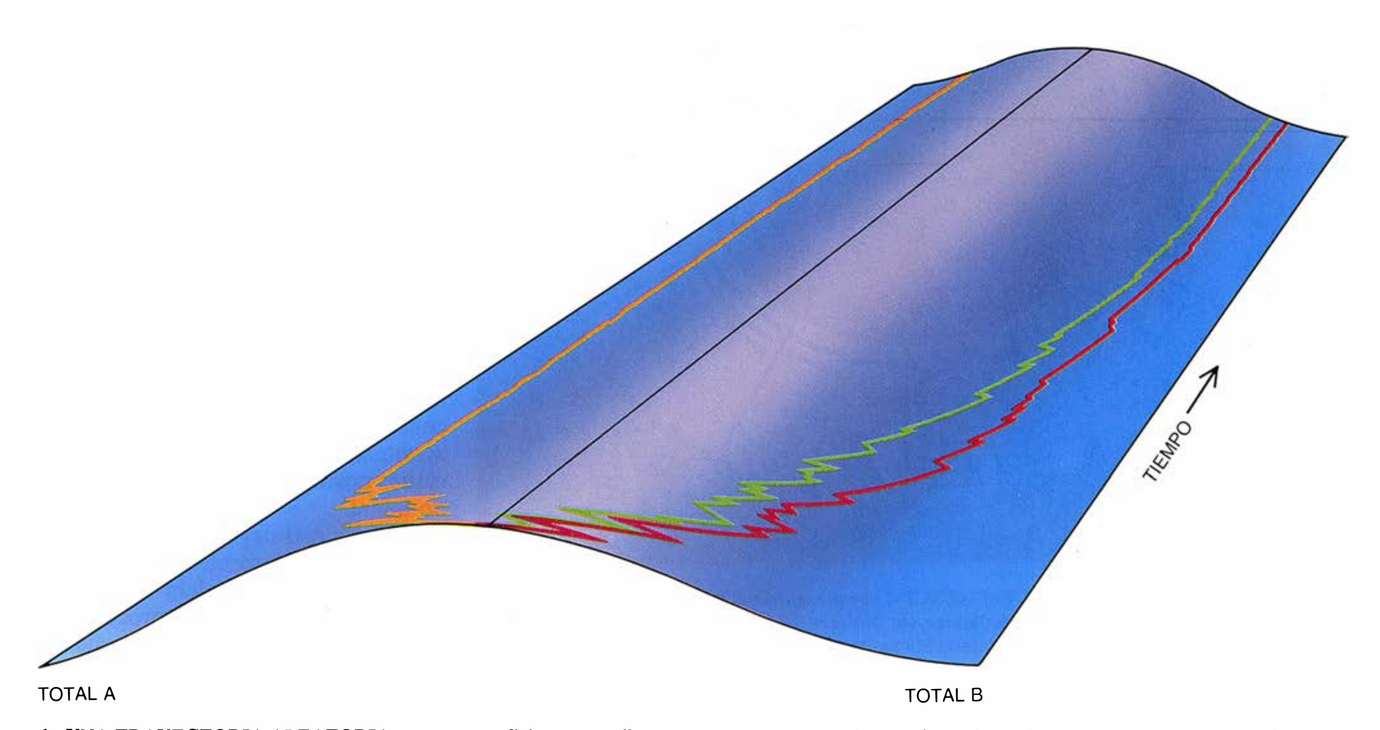
La evolución del mercado del vídeo no habría sorprendido al gran economista de la época victoriana Alfred Marshall, uno de los fundadores de la economía convencional de hoy en día. En sus *Principios de economía*, escritos en 1890, advirtió que si bien, por lo general, los costos de producción de una firma disminuyen cuando su parte del mercado aumenta, toda empresa que por simple suerte consiga al principio una elevada proporción del mercado estará en condiciones de aventajar a sus rivales; "toda firma que arranque con buen pie" acapararía el mercado. Marshall, sin embargo, no desarrolló esa observación y la teoría económica la ha ignorado casi siempre hasta hace poco.

Marshall no creía que los rendimientos crecientes se aplicasen por doquier; la agricultura y la minería –pilares de la economía de su tiempo– estaban sometidas a los rendimientos decrecientes causados por las limitadas disponibilidades de tierra fértil y de yacimientos minerales de alta calidad. La industria, en cambio, gozaba de unos rendimientos crecientes porque la gran fábrica permitía una mejor organización. Los economistas modernos, sin embargo, no piensan que las economías de escala conduzcan siempre a unos rendimientos crecientes. Las grandes instalaciones han resultado a veces más económicas; en muchas otras ocasiones, eso no ha sido así.

Me voy a permitir actualizar la reflexión de Marshall señalando que aquellas partes de la economía que se basan en recursos (agricultura, producción de bienes al por mayor, minería) siguen estando, en su mayoría, sometidas a rendimientos decrecientes. Aquí, la economía convencional continúa siendo válida. Las partes de la economía que se basan en conocimientos, en cambio, registran en su mayoría rendimientos crecientes. Bienes tales como ordenadores y sus soportes lógicos, productos farmacéuticos, misiles, aeronaves, automóviles, equipo de telecomunicaciones o artículos de fibra óptica son complicados de diseñar y fabricar. Requieren grandes inversiones iniciales en investigación, desarrollo y maquinaria, pero, una vez que empiezan las ventas, la producción marginal resulta relativamente barata. Diseñar, desarrollar, obtener el visto bueno de las autoridades y empezar a producir un nuevo fuselaje o un nuevo motor de avión, por ejemplo, suele costar de 2000 a 3000 millones de dólares. Después, cada ejemplar costará, quizás, entre 50 y 100 millones de dólares. Conforme se produzcan más unidades, los costos unitarios seguirán bajando y aumentarán los beneficios.

La mayor producción reportará ventajas adicionales: fabricar más unidades significa adquirir más experiencia del proceso de producción y conocer mejor cómo producirlas aún más baratas. Además, la experiencia obtenida con un bien o tecnología facilitará fabricar nuevos productos que incorporen tecnologías similares o conexas. El Japón, por ejemplo, convirtió una inversión inicial en instrumentos de precisión en capacidad para fabricar productos electrónicos de consumo y, a continuación, los circuitos integrados que incorporan.

No sólo los costos de producir bienes de alta tecnología disminuyen cuando una empresa los fabrica en mayor cantidad: también las ventajas que entraña su uso aumentan. Artículos como ordenadores o equipo de telecomunicación funcionan en redes que requieren compatibilidad; cuando una marca se hace con una parte apreciable del mercado, la gente tendrá muchas razones para comprar más del mismo producto, a fin de




**1. UNA TRAYECTORIA ALEATORIA en una superficie convexa ilustra la competencia con rendimientos crecientes entre dos tecnologías. El azar determina las primeras pautas de adopción e influye así en la rapidez con que cada competidor mejora. A medida que una tecnología adquiera más seguidores (que corresponde al movimiento descendente hacia un borde de la superficie), resultará cada vez más probable su adopción ulterior.**

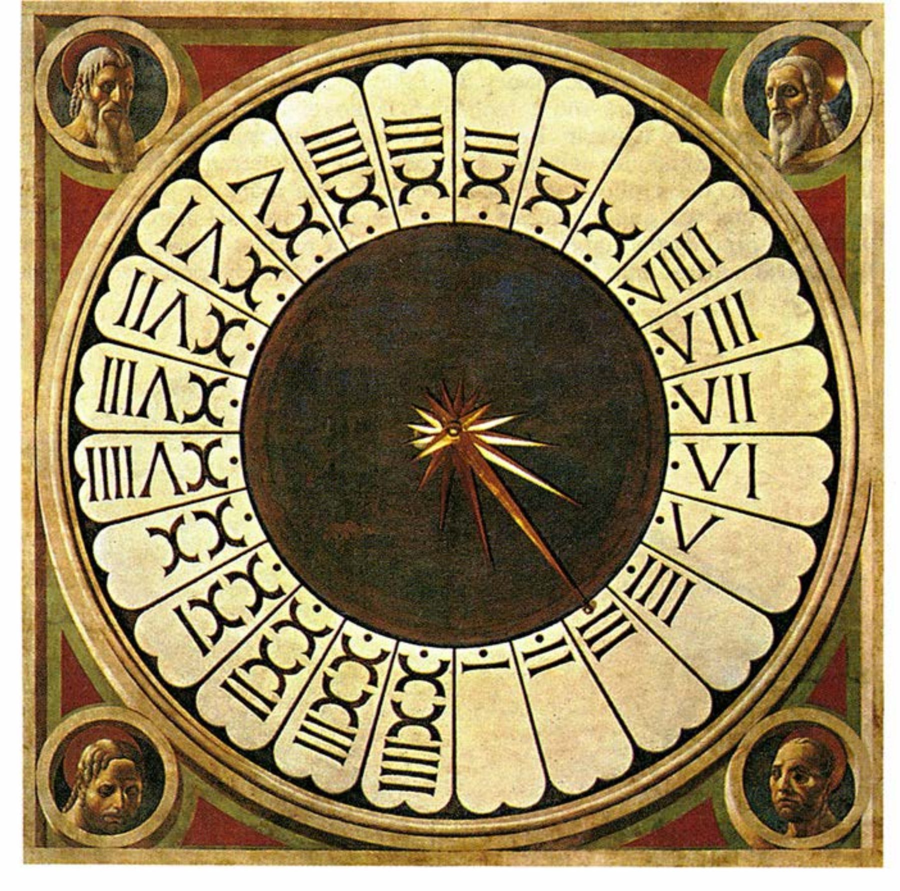
poder intercambiar información con quienes ya lo usan.

Si los rendimientos crecientes son importantes, ¿por qué han sido casi siempre ignorados hasta hace poco? Algunos dirían que los productos complicados, o la alta tecnología, donde intervienen con tanto peso los rendimientos crecientes, constituyen un acontecimiento reciente. Es verdad, pero tal cosa sólo constituye una respuesta parcial. Después de todo, en los decenios de 1940 y 1950, los economistas Gunnar K. Myrdal y Nicholas Kaldor, entre otros, identificaron mecanismos de retroacción positiva que no se veían afectados por la tecnología. Los economistas ortodoxos han evitado los rendimientos crecientes por motivos más profundos.

Algunos economistas veían con disgusto, e incluso consideraban poco científico, el hecho de que existiera más de una solución a un mismo problema. "Los equilibrios múltiples", escribía Joseph A. Schumpeter en 1954, "no son por fuerza inútiles, pero desde el punto de vista de *toda* ciencia exacta la existencia de un equilibrio determinado unívocamente reviste, claro está, suma importancia, aun cuando ese resultado haya de obtenerse a costa de supuestos muy restrictivos; si no hay posibilidad de demostrar la existencia de un equilibrio unívocamente determinado –o, como mucho, de un pequeño número de posibles equilibrios– aunque sea en un nivel de abstracción muy elevado, el conjunto de los acontecimientos de que se trate se vuelve un caos que escapa a todo control analítico."

Otros economistas se percataban de que las teorías que incorporaban los rendimientos crecientes destruirían su mundo familiar de equilibrios únicos y predecibles y la idea de que la elección del mercado era siempre la mejor. Asimismo, si una empresa o unas pocas empresas llegaban a dominar un mercado, el supuesto de que ninguna firma se bastaba por sí sola para condicionar los precios del mercado (lo que hace que los problemas económicos sean fáciles de analizar) también se vendrían abajo. Cuando John R. Hicks estudió esas posibilidades en 1939, se echó atrás alarmado. "El edificio amenazado", escribió, "es el constituido por la mayor parte de la teoría económica". Los economistas se limitaron así a los rendimientos decrecientes, que no presentaban anomalías y podían analizarse cabalmente.

Sin embargo, otros manifestaron su perplejidad ante la pregunta de cómo un mercado podía seleccionar una entre varias soluciones posibles. En el ejemplo de Marshall, la firma que es la más grande al comienzo tiene los costos de producción más bajos y ha de triunfar inevitablemente en el mercado. En tal caso, ¿por qué las empresas menores se molestarían en competir? Por otra parte, si por ventura un mercado comenzara con varias firmas idénticas, sus cuotas de mercado se mantendrían en un equilibrio inestable por siempre jamás.



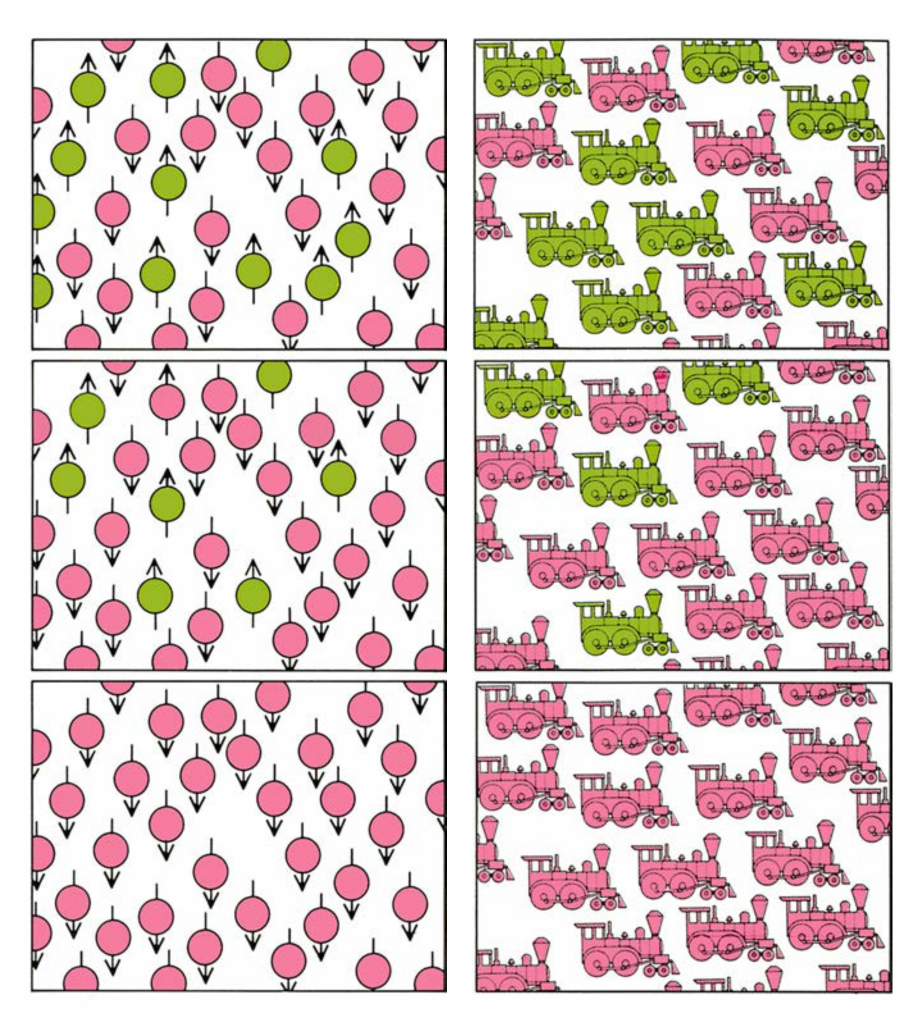

**2. RELOJ DE LA CATEDRAL DE FLORENCIA, cuyas manecillas se mueven "en sentido antihorario", recorriendo una esfera donde figuran las 24 horas del día. Cuando Pablo Uccello diseñó el reloj en 1443, todavía no se había establecido convención alguna para las esferas de los relojes. Los diseños que competían entre sí estaban sometidos a los rendimientos crecientes: cuantas más esferas de reloj de un tipo se fabricaran, más personas acabarían acostumbrándose a leerlas. Por consiguiente, lo más probable era que las esferas de los relojes futuros fueran todas de la misma índole. Después de 1550, los diseños "en el sentido de las manecillas del reloj" que sólo marcaban 12 horas habían desplazado a todos los demás. El autor del artículo sostiene que acontecimientos fortuitos unidos a una retroacción positiva determinarán muchas veces los acontecimientos económicos.**

Al estudiar estos problemas en 1979, pensé que podía encontrar una salida a muchas de tales dificultades. En el mundo real, si varias firmas de tamaño similar entran en un mercado al mismo tiempo, pequeños acontecimientos fortuitos –pedidos imprevistos, encuentros afortunados con compradores, genialidades o fantasías empresariales– contribuirán a determinar quiénes se hacían con las primeras ventas y qué firma dominaba con el tiempo. La actividad económica es el resultado de agregar transacciones individuales que por sus exiguas dimensiones no se pueden observar, y esos pequeños acontecimientos "aleatorios" pueden acumularse y verse ampliados por retroacciones positivas que acabarán determinando el resultado. Esos hecho sugerían que situaciones dominadas por los rendimientos crecientes no deberían contemplarse como problemas estáticos y deterministas, sino como procesos dinámicos basados en acontecimientos fortuitos y retroacciones positivas naturales, de carácter no lineal.

Con esa estrategia, un mercado de rendimientos crecientes podría reconvertirse en un modelo teórico y observarse conforme el proceso correspondiente se desarrolla más y más. En ocasiones surgiría una solución, otras veces (en idénticas condiciones) una distinta. Sería imposible saber de antemano cuál de las muchas soluciones se impondría en el discu-

rrir económico. Con todo, sería posible registrar el conjunto de acontecimientos aleatorios que conducen a cada solución y estudiar la probabilidad de que surgiera una solución dada bajo un determinado conjunto de condiciones iniciales. La idea era sencilla, y pudo muy bien habérsele ocurrido a economistas del pasado. Pero llevarla a la práctica requería la teoría de los procesos aleatorios no lineales, inexistente en aquel entonces.

Ningún problema de rendimientos crecientes ha de estudiarse de modo aislado; muchos acaban encajando en un esquema probalístico general de carácter no lineal. Tal cosa puede describirse si imaginamos una mesa sobre la que colocamos bolas de una en una; esas bolas pueden ser de varios colores, a saber, blancas, rojas, verdes o azules. Desconocemos el color de la bola que vamos a añadir a las que ya estén sobre la mesa, pero la probabilidad de un determinado color dependerá de la proporción de colores que haya en la mesa. Si una proporción creciente de bolas de un determinado color aumenta la probabilidad de que se añada otra bola de igual color, el sistema presenta una retroacción positiva. La cuestión es cómo saber, dada la función que relacione las proporciones existentes con las probabilidades, cuáles serán las proporciones de cada color en la mesa después de que se hayan añadido muchas bolas.



**3. FERROIMANES Y ANCHOS DE VIA REGIONALES DE LOS FERROCARRILES acaban ordenándose del mismo modo. Cuando un material magnético en desorden se enfría (*izquierda*), los dipolos atómicos que contiene ejercen fuerzas recíprocas que hacen que los dipolos próximos apunten en la misma dirección. Llega un momento en que todos los dipolos de una muestra apuntan en la misma dirección, pero el sentido que adopten (hacia arriba o hacia abajo) no puede predecirse con antelación. De igual modo, tal como Douglas Puffert, del Colegio Swarthmore, ha mostrado, los ferrocarriles privados colindantes (*derecha*) del siglo pasado adoptaron el mismo ancho para ampliar su alcance más fácilmente. Andando el tiempo, ocurrió que todos (o casi todos) los ferrocarriles usaban el mismo ancho. Ecuaciones similares describen el comportamiento de ambos sistemas.**

En 1931, el matemático George Polya resolvió una versión muy particular de ese problema donde la probabilidad de añadir un color era siempre igual a la proporción que había de ese color sobre la mesa. Tres teóricos estadounidenses del cálculo de probabilidades, Bruce M. Hill, de Ann Arbor, y David A. Lane y William D. Sudderth, de la Universidad de Minnesota, resolvieron una versión no lineal más general en 1980. En 1983, dos teóricos soviéticos de las probabilidades, Yuri M. Ermoliev y Yuri M. Kaniovski, ambos del Instituto Glushkov de Cibernética de Kiev, y yo mismo, encontramos la solución de una versión muy general. Demostramos que, conforme sigan añadiéndose bolas, la proporción de cada color acabará estabilizándose en un "punto fijo" de la función de probabilidad, lo que arroja un conjunto de valores donde la probabilidad de añadir cada uno de los colores es igual a la proporción de ese color sobre la mesa. Los rendimientos crecientes permiten varios conjuntos de puntos fijos de esa índole.

Esto significa que podemos determinar las pautas o soluciones posibles de un problema de rendimientos crecientes resolviendo la cuestión mucho más sencilla de encontrar los conjuntos de puntos fijos de su función de probabilidad. Con tales instrumentos, los economistas pueden ahora definir los problemas de rendimientos crecientes con precisión, identificar sus soluciones posibles y estudiar el proceso por el que se llega a una solución. Los rendimientos crecientes ya no son "un caos que escapa a todo control analítico".

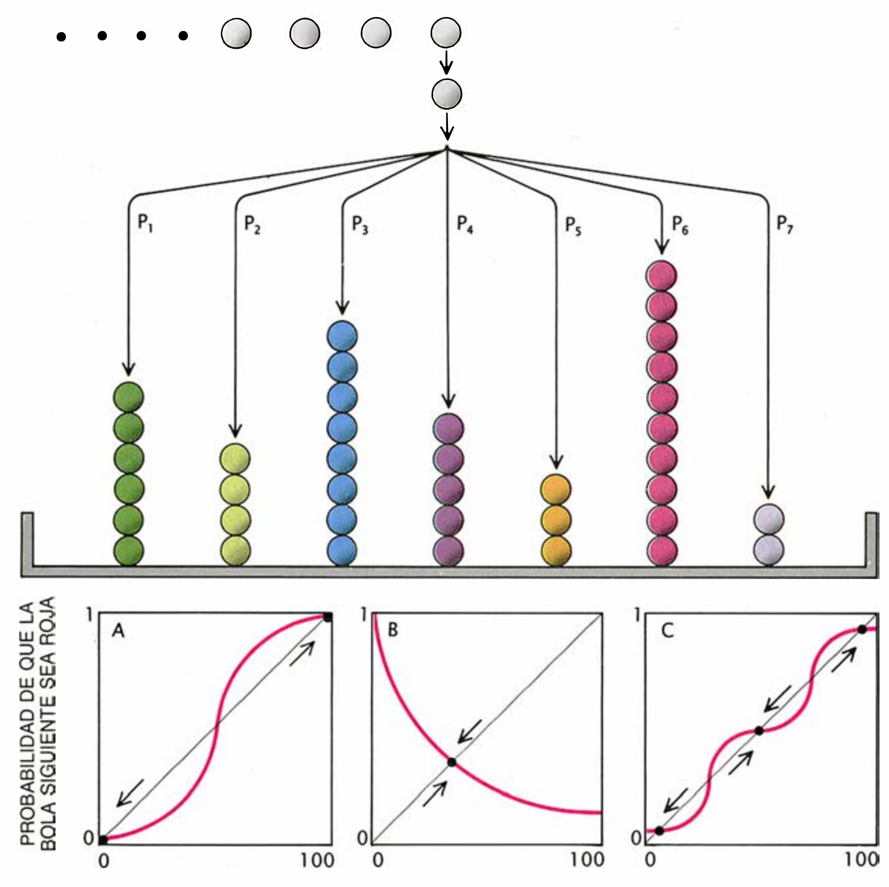
En el mundo real, las bolas podrían ser empresas; sus colores, las regiones donde decidan establecerse. Supongamos que las empresas entran en un sector una por una y escogen su ubicación con miras a maximizar el beneficio. La preferencia geográfica de cada empresa (el beneficio intrínseco que obtiene al establecerse en una región particular) variará; la suerte determina la preferencia de la empresa que entre a continuación en ese sector. Supongamos también, sin embargo, que los beneficios de una empresa aumentan cuando se halla cerca de otras empresas (que son proveedores o clientes suyos). La primera empresa que entra en el sector elige un emplazamiento basado únicamente en preferencias geográficas. La segunda empresa decide basándose en la preferencia modificada por los beneficios que obtendrá al ubicarse cerca de la primera empresa. La tercera empresa se ve influida por las posiciones de las dos primeras, y así sucesivamente. Si alguna ubicación tiene la buena fortuna de atraer un mayor número de empresas que las

demás en las primeras etapas de esa evolución, la probabilidad de que atraiga más empresas aumenta. La concentración industrial se refuerza a sí misma.

La secuencia histórica aleatoria de empresas que entran en el sector determinará la pauta de las ubicaciones regionales resultantes, pero la teoría muestra que no todas las pautas son posibles. Si la atracción ejercida por la presencia de otras empresas aumenta continuamente cuando se añaden más empresas, alguna región siempre predominará y eliminará a todas las demás. Si la atracción no aumenta, otras soluciones, con una distribución de la actividad entre varias regiones, serán posibles. Nuestros nuevos instrumentos nos dicen qué tipos de solución pueden darse en cada caso.

¿Hay regiones que cuentan, en la práctica, con una gran proporción de una industria debido al azar histórico, más que a su superioridad geográfica? El condado de Santa Clara en California ("Valle del Silicio") es un ejemplo que viene a punto. En los años cuarenta y primeros cincuenta, ciertas personas que ocupaban puestos claves en la industria electrónica de los Estados Unidos –los hermanos Varian, William Hewlett y David Packard, William Shockley– se establecieron cerca de la Universidad de Stanford; la disponibilidad local de ingenieros, suministros y componentes que esas primeras empresas ayudaron a crear hicieron al condado de Santa Clara extremadamente atractivo para las 900 empresas más o menos que vinieron después. Si esos primeros empresarios hubieran preferido otros lugares, la concentración más densa de electrónica del país podría muy bien estar ubicada en otra parte.

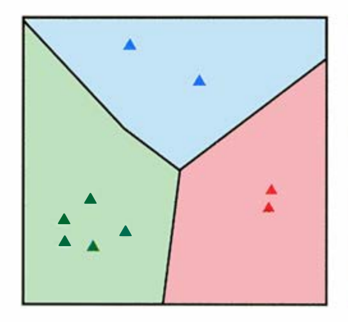
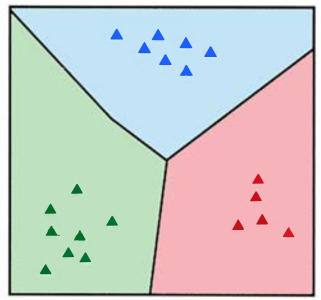
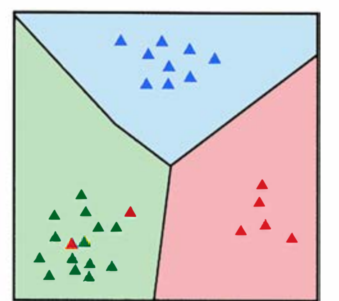
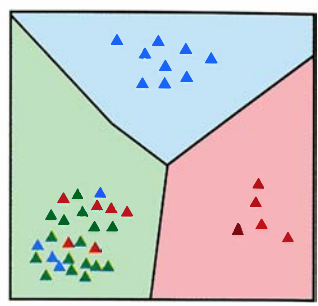
En una escala mayor, cabe preguntarse si, en el caso de que pequeños acontecimientos históricos hubieran sido diferentes, el emplazamiento de las ciudades mismas no habría sido distinto. Pienso que la respuesta es positiva. En tanto en cuanto ciertas ubicaciones constituyen puertos naturales o encrucijadas al borde de ríos o lagos, la distribución de las ciudades de hoy no obedece al azar sino a la geografía. En tanto en cuanto las actividades económicas y la gente se ven atraídas por lugares donde esos recursos ya están reunidos, unas pequeñas y primeras concentraciones fortuitas pueden haber sido la semilla de la configuración actual de los centros urbanos. "El azar y la necesidad", para usar una frase de Jacques Monod, se influyen recíprocamente. Ambos aspectos han desempeñado papeles cruciales en el desarrollo de las ciudades en los Estados Unidos y por doquier.




**4. TEORIA NO LINEAL DE PROBABILIDADES, que predice el comportamiento de sistemas sometidos a rendimientos crecientes. En este modelo, bolas de diferentes colores se van colocando sobre una mesa; la probabiidad de que la bola siguiente sea de un determinado color dependerá de las proporciones de colores existentes (*arriba*). Los rendimientos crecientes se dan en *A* (las curvas muestran el caso de dos colores; las flechas indican sentidos de dirección probables): una bola roja tendrá más probabilidad de añadirse cuando ya exista una elevada proporción de bolas de ese color. Ese caso presenta dos puntos de equilibrio: uno en el que casi todas las bolas son rojas; otro, en el que muy pocas son rojas. En *B* hay unos rendimientos decrecientes, esto es, una proporción mayor de bolas rojas disminuye la probabilidad de que se añada otra. Se trata de un punto de equilibrio singular. La combinación de rendimientos crecientes y decrecientes (*C*) arroja muchos puntos de equilibrio.**

Los mecanismos que se refuerzan a sí mismos, distintos de los regionales, operan en la fabricación y en el comercio internacionales relacionados con la alta tecnología. Los países que adquieran mucha importancia y experiencia en un sector de tecnología elevada podrán obtener ventajas de menores costos y mayor calidad que les permitan excluir a otros países. Por ejemplo, en los primeros años de la década de 1970, los fabricantes japoneses de automóviles empezaron a vender cantidades apreciables de coches pequeños en los Estados Unidos. Conforme el Japón aumentaba su parte del mercado sin mucha oposición de Detroit, sus ingenieros y trabajadores adquirieron experiencia, sus costes disminuyeron y sus productos mejoraron. Tales factores, unidos a unas mayores redes de venta, permitieron al Japón aumentar su parte en el mercado estadounidense; como consecuencia de ello, los trabajadores adquirieron aún más experiencia, los costes bajaron todavía más y la calidad siguió mejorando. Antes de que Detroit respondiera seriamente, ese proceso de retroacción positiva había permitido a las compañías japonesas hacerse con una parte importante del mercado de automóviles pequeños de los Estados Unidos. Procesos similares han tenido lugar en los mercados de aparatos de televisión, circuitos integrados y otros productos.

¿Cómo deberían los países responder ante una economía mundial donde se aplican tales normas? Las recomendaciones convencionales de política comercial basadas en rendimientos constantes o decrecientes tienden a adoptar enfoques pruden-

**5. LAS EMPRESAS ELIGEN SU UBICACION para maximizar los beneficios, que vienen determinados por una preferencia geográfica intrínseca (indicada por un color) y por la presencia de otras empresas. En este ejemplo, generado por ordenador, casi todas las primeras empresas se establecen en la región verde, y así todas las compañías de nueva creación acaban instalándose allí. Semejante apiñamiento parecería implicar que la región verde es algo superior. En otras aplicaciones del programa, sin embargo, dominan las regiones de color rojo y azul. (Dibujos de Paul Wallich.)**

tes. Se basan en el mercado abierto, desalientan los monopolios y dejan asuntos tales como el gasto en investigación y desarrollo en manos de las empresas. Su hipótesis básica es que existe un precio mundial fijo al que los productores lanzan sus bienes al mercado; por tanto, toda interferencia con los costos y precios locales mediante subvenciones o aranceles resulta poco productiva. Tales políticas son apropiadas para las partes de rendimientos decrecientes de la economía, pero no para las partes basadas en la tecnología, donde predominan los rendimientos crecientes.

Las políticas que resultan apropiadas para tener éxito en la producción y el comercio internacional de alta tecnología incitarían a las industrias a que procurasen de un modo agresivo mejorar productos y procesos. Reforzarían la base nacional de investigación en la que se asientan las ventajas de la alta tecnología. Fomentarían que las empresas de un determinado sector pusieran en común sus recursos en empresas mixtas que se repartieran costos variables, redes de comercialización, conocimientos técnicos y normas de fabricación. Cabría incluso alentar alianzas estratégicas, permitiendo que compañías de varios países entrasen en una industria compleja que ninguna por sí sola podría abordar. La teoría de los rendimientos crecientes también destaca la importancia de elegir el momento oportuno al emprender investigaciones en nuevas industrias. No tiene mucho sentido entrar en un mercado próximo a cerrarse o que, de otro modo, brinde pocas posibilidades de éxito. Políticas de esa índole están poco a poco propugnándose y adoptándose en los Estados Unidos.

El valor de otras políticas, tales como subvencionar y proteger industrias nuevas –la bioingeniería, por ejemplo– para capturar mercados extranjeros, resulta discutible. A veces se han citado dudosas ventajas retroactivas para justificar actividades inútiles patrocinadas por los gobiernos. Asimismo, tal como Paul R. Krugman, del Instituto de Tecnología de Massachusetts, y varios otros economistas han señalado, si un país aplica tales políticas, otros tomarán medidas de represalia subvencionando sus propias industrias de alta tecnología. Nadie gana con ello. La cuestión de una política industrial y comercial óptica basada en unos rendimientos crecientes ya se está estudiando intensivamente. Las políticas que los países elijan determinarán no sólo lo que será la economía mundial en el decenio de 1990, sino también quiénes ganarán y quiénes perderán.

Los mecanismos de rendimientos crecientes no se limitan a decidir las balanzas competitivas entre las naciones; pueden provocar también que las economías –aun cuando hayan tenido tanto éxito como las de los Estados Unidos y el Japón– se queden encerradas en unas trayectorias de desarrollo menos buenas. Una tecnología que mejore lentamente al principio, pero con muchas posibilidades a la larga, podría fácilmente anularse, abocando a la economía a una vía que sea a la vez inferior y difícil de evitar.

Las tecnologías suelen mejorar cuando las adoptan más gente y las empresas adquieren una experiencia que orienta el desarrollo ulterior. Ese vínculo constituye un proceso retroactivo positivo: cuanto más gente adopte una tecnología, más mejorará y más atractiva resultará. Cuando dos o más tecnologías (o dos o más productos) compiten, la retroacción positiva hace que su mercado resulte inestable. Si una de ellas se destaca en el mercado, quizá por azar, su desarrollo podrá acelerarse lo bastante como para apropiarse del mercado. Una tecnología que mejore más rápidamente conforme más personas la adopten tendrá más posibilidades de sobrevivir, ya que contará con una "ventaja selectiva". La superioridad inicial, sin embargo, no garantiza la idoneidad a largo plazo.

En 1956, por ejemplo, cuando los Estados Unidos emprendieron su programa de energía nuclear, se propusieron varios diseños: reactores enfriados por gas, agua ligera, agua pesada, incluso sodio líquido. Robin Cowan, de la Universidad de Nueva York, ha mostrado que una serie de circunstancias triviales empujaron prácticamente a toda la industria nuclear de los Estados Unidos hacia el agua ligera. Los reactores que la usaban se adaptaron inicialmente a partir de unidades compactas, diseñadas para impulsar submarinos nucleares. El papel de la Marina estadounidense en los primeros contratos de construcción de reactores, los esfuerzos desplegados por el Consejo de Seguridad Nacional para obtener un reactor –el que fuere– que operara en tierra, a raíz del lanzamiento del *Sputnik* en 1957, así como las predilecciones de algunos funcionarios importantes, todo ello actuó en favor del predominio inicial de los reactores de agua ligera. La experiencia en la construcción permitió mejorar los diseños correspondientes y, hacia 1965, la trayectoria de la industria quedaba fijada. El que otros diseños hubieran sido en realidad superiores a la larga es algo discutible, pero gran parte de las obras técnicas que se han escrito sobre el particular sugiere que unos reactores de alta temperatura, enfriados por gas, habrían sido mejores.

Las convenciones o normas técnicas, así como determinadas tecnologías, tienden a quedarse bloqueadas por una retroacción positiva, como mi colega Paul A. David, de Stanford, ha documentado para varios casos históricos. Aunque una norma por sí sola puede no mejorar con el tiempo,

su adopción con carácter general hace ventajoso para quienes llegan por vez primera a una actividad –y que han de intercambiar información o productos con quienes ya trabajan en ella– ceñirse a lo ya establecido, trátese del idioma inglés, un sistema de televisión de alta definición, el sentido en que ha de girar una rosca o el teclado de una máquina de escribir. Normas que se han establecido pronto –tales como el lenguaje de programación FORTRAN de los años cincuenta– pueden ser de difícil sustitución por otras, por muy superiores que éstas sean.

Hasta hace poco, los textos económicos convencionales tendían a describir la economía como algo parecido a un gran sistema newtoniano, con una única solución de equilibrio, predeterminada por aspectos tales como recursos minerales, geografía, población, gustos del consumidor y posibilidades tecnológicas. Desde ese punto de vista, las perturbaciones o alteraciones temporales –tales como la crisis del petróleo de 1973 o el hundimiento del mercado bursátil en 1987– se ven rápidamente compensadas por las fuerzas que suscitan. Dadas las posibilidades tecnológicas futuras, se podría, en teoría, prever con precisión la trayectoria de la economía como una solución, con suaves cambios, de las ecuaciones analíticas que rigen precios y cantidades de bienes. La historia, en esta perspectiva, no resulta demasiado importante; simplemente conduce la economía a su inevitable equilibrio.

La economía de retroacciones positivas, en cambio, encuentra su parangón en la física no lineal moderna. Los materiales ferromagnéticos, los vidrios espinoriales, los láseres de estado sólido y otros sistemas físicos que consisten en elementos que se refuerzan mutuamente muestran las mismas propiedades que los ejemplos económicos que he citado. Presentan una "fijación de fase" en una de las muchas configuraciones posibles; pequeñas perturbaciones en momentos críticos influyen en el resultado que se elegirá, y el resultado escogido puede tener mayor energía (esto es, ser menos favorable) que otros posibles estados físicos.

Esta clase de economía también tiene parecido con la teoría evolutiva del equilibrio por puntos. Los pequeños acontecimientos (las mutaciones de la historia) suelen ser ignorados, pero de vez en cuando adquieren suma importancia, al inclinar parte de la economía hacia estructuras y pautas nuevas que, retenidas entonces, constituyen una nueva etapa del desarrollo.

En esa nueva perspectiva, economías inicialmente idénticas con sectores apreciables de rendimientos crecientes no elegirán necesariamente la misma trayectoria. Al contrario, acabarán divergiendo. En la medida en que los pequeños sucesos que determinan la trayectoria general quedan fuera del alcance de la lupa de los economistas, una previsión precisa del futuro de la economía puede resultar, no sólo en la práctica, sino también en la teoría, inviable. Dirigir una economía con retroacciones positivas hacia el mejor de sus muchos estados de equilibrio posibles requiere mucha suerte y elegir bien el momento, esto es, apreciar cuándo hay más posibilidades de un cambio beneficioso de una pauta a otra. La teoría puede ayudar a que se identifiquen esos estados y momentos y orientar a quienes deciden para que apliquen la cantidad adecuada de esfuerzo (sin pecar por defecto ni por exceso), a fin de cambiar estructuras cerradas en sí mismas.

El filósofo inglés de la ciencia Jacob Bronowski señaló en cierta ocasión que la economía ha padecido desde hace tiempo una estructura fatalmente simple impuesta en el siglo XVIII. Por mi parte, encuentro apasionante que tal cosa esté ahora cambiando. Con la aceptación de las retroacciones positivas, la teoría económica está empezando a definir la economía no como algo sencillo sino complejo, no como una ciencia determinista, predecible y mecanicista, sino como un conocimiento orgánico, vinculado a procesos y en constante evolución.


BIBLIOGRAFIA COMPLEMENTARIA

MARKET STRUCTURE AND FOREING TRADE. Elhanan Helpman y Paul Krugman. The MIT Press, 1985.

PATH-DEPENDENT PROCESSES AND THE EMERGENCE OF MACRO-STRUCTURE. W. Brian Arthur, Yu M. Ermoliev y Yu M. Kaniovski en *European Journal of Operational Research*, vol. 30, págs. 294-303; 1987.

SELF-REINFORCING MECHANISMS IN ECONOMICS. W. Brian Arthur en *The Economy as an Evolving Complex System*. Dirigido por Philip W. Anderson, Kenneth J. Arrow y David Pines. Addison-Wesley Publishing Co., 1988.

PATH-DEPENDENCE: PUTTING THE PAST INTO THE FUTURE OF ECONOMICS. Paul David. I.M.S.S.S. Tech Report N.° 533, Stanford University; noviembre de 1988.

COMPETING TECHNOLOGIES, INCREASING RETURNS, AND LOCK-IN BY HISTORICAL EVENTS. W. Brian Arthur en *The Economic Journal*, vol. 99, n.° 394, págs. 116-131; marzo de 1989.

# Taller y laboratorio

*Cuando se estira una hoja de polímero, ésta sufre un adelgazamiento local antes de desgarrarse*

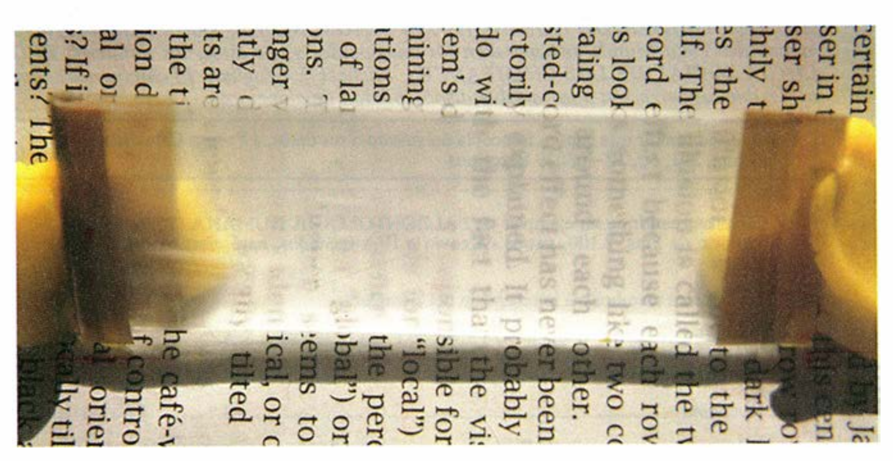
Jearl Walker

Muchos son los polímeros laminados, entre ellos los que se emplean para envolver alimentos, que se comportan de manera peculiar frente al estiramiento. Tales láminas, en efecto, ni se rompen de golpe como los hilos ni se estiran como las gomas. Al principio se resisten fuertemente al estiramiento, cediendo luego de repente con una disminución del espesor o de la anchura (perpendicularmente a la dirección de estiramiento), o de ambos. Tal estrechamiento suele conocerse como formación de un cuello, estricción o estirado en frío. (Esta tercera expresión se emplea por oposición a un estrechamiento parecido que sufren las barras de metal calientes cuando se estira de ellas.)

Probemos con una hoja de plástico elástico tirando de sus dos extremos. Al principio tendremos que tirar con fuerza para conseguir moverla algo, pero después el plástico se estira estrechándose repentinamente. Una vez llegados a este punto, el plástico presenta menos resistencia al estiramiento y la zona estrangulada comienza a propagarse hacia los extremos. Ahora bien, cuando esa propagación alcanza un grado apreciable, hay que volver a estirar con fuerza hasta que el plástico termina por desgarrarse por la zona estrechada.

Antes de que les exponga la mecánica del estiramiento, permítanme que les hable de un efecto bastante desconcertante de este fenómeno. Algunas hojas de polímero son transparentes, aunque oscurecen la imagen. Así, cuando pretendemos leer una página impresa a través de una de ellas, hay que mantener el plástico muy cerca del papel; conforme lo vamos alejando del texto, las letras no tardan en oscurecerse tanto que resultan ilegibles. Esa fosquedad obedece, sobre todo, a la difusión de la luz en las moléculas del plástico. La distancia a la cual es posible leer a través de la hoja constituye una medida de la intensidad con que se difunde la luz.

Si estiramos ahora la hoja de plástico de manera que sufra una estricción, la luz atravesará menos materia (o sea, menos moléculas) y la difusión resultante debería ser menor; así, deberíamos poder leer a través del plástico aun situándolo a una mayor distancia de las letras. Por lógico que parezca este razonamiento, es evidente que no concuerda con la situación que se presenta en la figura 1. Esa fotografía la obtuve estirando y provocando la estricción en la mitad de una banda de polímero, que luego tendí sobre un portaobjetos de microscopio. El portaobjetos lo coloqué inclinado, fijándolo con plastilina, sobre una página impresa, de tal modo que la distancia entre ésta y la banda fuera variable. Allí donde la zona no afectada por la estricción se encuentra más alejada de la página, el texto se ve borroso aunque legible, mientras que en la parte estrangulada el texto se percibe borroso del todo, incluso donde la banda está muy cerca del papel. La causa es sutil; lo veremos más adelante.



*1. Oscurecimiento de la letra impresa con una tira de polímero sometida a estricción.*

Un polímero es una molécula de gran tamaño, formada por la repetición de cierta unidad química llamada monómero. La gran cantidad de polímeros diferentes con que podemos toparnos en nuestra vida cotidiana presenta una amplísima gama de propiedades mecánicas y ópticas. De entre los que exhiben estricción, he elegido dos para tratar aquí de ellos. Uno es el polietileno, que suele encontrarse en las cocinas en forma de saquitos, bolsitas u hojas (a menudo, autoadhesivas) para guardar o envolver alimentos. El monómero del polietileno es una sencilla agrupación de dos átomos de carbono y cuatro de hidrógeno. En una lámina de polietileno, las zonas donde las grandes moléculas de polímero se agrupan en minúsculos cristales están separadas por zonas amorfas carentes por completo de organización. El otro polímero que he elegido es el parafilm, cuyas láminas encontraremos en cualquier laboratorio químico o biológico, donde se emplean para sellar vasos y otros recipientes. El parafilm es una mezcla de parafina y polietileno.

La estricción es en gran medida el resultado de la orientación de las moléculas del polímero. Consideremos una hoja de polímero cuyo espesor mengüe, e imaginemos que pudiéramos ver sus moléculas. Inicialmente podría ser que estuvieran organizadas a pequeña escala (cristalizadas en parte, o más o menos alineadas por el proceso de fabricación); o puede incluso que estuvieran orientadas completamente al azar. Cuando tiramos, la hoja sólo podrá alargarse si las moléculas son capaces de girar o desplazarse alineándose con la dirección de la fuerza y adaptarse así al aumento de longitud. Al principio, los enlaces químicos se oponen a la reorientación: la hoja sólo se estira un poco. Pero en cuanto la fuerza alcanza cierto valor crítico, las moléculas de la zona más débil de la hoja ceden, rompen sus enlaces más débiles, se deslizan unas sobre otras y tienden a alinearse [*figura 2*]. En esa zona, la hoja de plástico cede a regañadientes, sacrificando su espesor a la longitud. Aunque, por otra parte, gana resis-

tencia mecánica, en el sentido de que las moléculas reorientadas han roto sus uniones más débiles y se quedan con las más fuertes.

Si seguimos tirando, se alinearán más moléculas en la zona de estricción y disminuirá aún más el espesor. Cuando ya estén alineadas un gran número de moléculas, el espesor se habrá reducido al máximo; los enlaces entre las moléculas serán, entonces, demasiado fuertes para permitir que prosiga la deformación. Si no dejamos de tirar, las moléculas que a continuación se rindan serán las situadas en los "cuellos" de la estricción, con lo cual ésta se propagará en la dirección de la fuerza aplicada. Cuando la estricción se haya extendido a toda la hoja, la naturaleza "endurecida" de las interconexiones de las moléculas, que afectará ya a toda la hoja, hará necesario que volvamos a tirar con fuerza para estirarla todavía más y no tardará en desgarrarse antes que seguir cediendo.

Para preparar el parafilm del que me serví en el experimento, empecé cortando una banda rectangular con unas tijeras, la coloqué bien plana sobre una mesa y luego le adosé a los extremos cinta adhesiva, de embalar, muy robusta. Con regla y pluma, tracé líneas a lo ancho de la banda a intervalos de dos milímetros. Después llevé la banda a un gato de laboratorio, apretando con fuerza los extremos encintados contra la cara del gato. (En lugar de éste, puede emplearse un tornillo de banco.) Así la banda quedó extendida verticalmente entre las caras del gato; lisa y plana, aunque no tensa

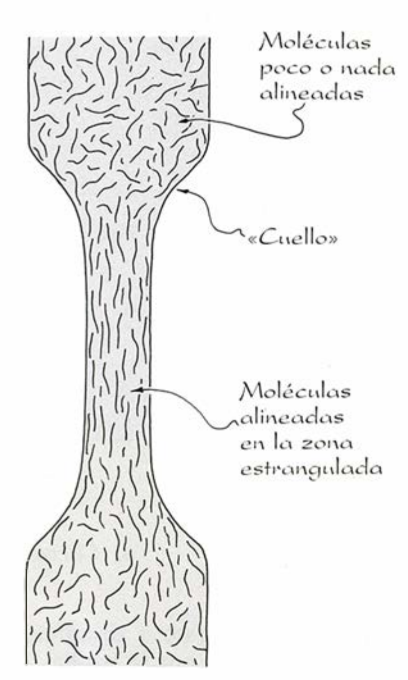
Medí y anoté la longitud vertical de la banda y su anchura de izquierda a derecha. Para medir el espesor con el micrómetro, aflojé primero el tope móvil del instrumento lo suficiente para permitir que la banda pasara por entre ambos topes y, después, apreté éstos gradualmente a la vez que movía despacio el micrómetro en pequeños círculos. Cuando los topes se encontraron lo bastante cercanos para apresar la banda de plástico, los aflojé lo justo para que no apresaran la tira; anoté en ese momento la lectura del micrómetro. Medición que repetí varias veces y saqué la media de los resultados. La banda presentaba unos 0,14 milímetros de espesor.

Entonces comencé a girar el tornillo del gato, separando las caras y estirando la banda. Al principio costaba hacer girar el tornillo, pero cuando hube estirado la banda del orden del 9 por ciento, cedió bruscamente y apareció un adelgazamiento localizado en una estrecha franja que atravesaba la tira de un lado a otro. Descubrí que las líneas de tinta estaban separadas un 25 por ciento más en la zona del estrangulamiento, mientras que las líneas del resto de la banda no mostraban ningún tipo de separación adicional.

Conforme proseguía girando el tornillo del gato, la estricción progresaba poco a poco hacia las dos caras del útil, alargando la separación de las líneas que encontraba en su avance. Este no era uniforme: tras cada vuelta del tornillo, en los "hombros" de la zona de estricción se notaban unas "islas" de material no estirado visiblemente distinto del material estirado que las rodeaba. A cada vuelta del tornillo, medía la longitud, la anchura y el espesor de la banda y la separación de las líneas en diversos puntos. Cuando hube alargado la banda un 90 por ciento aproximadamente, todas las líneas mostraban ya una separación adicional.

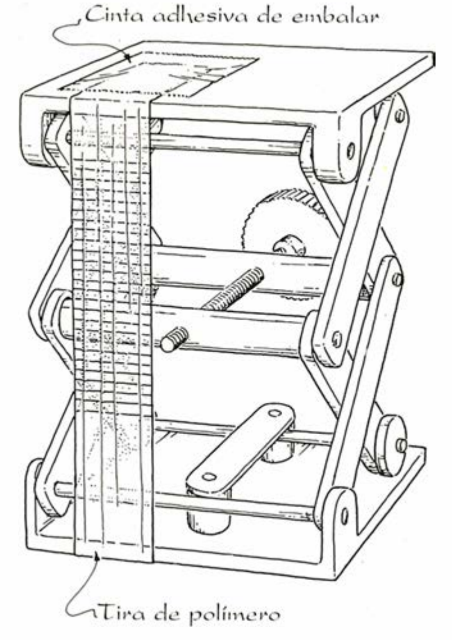
Alcanzado un alargamiento del 150 por ciento (o sea, cuando la tira se alargó hasta dos veces y media su longitud original), las líneas de la zona de estricción original estaban separadas hasta 9 milímetros (lo que corresponde a un alargamiento del 350 por ciento) y el espesor se había reducido a 0,064 milímetros. Las demás líneas estaban separadas 3,5 milímetros (el material se había estirado del orden del 75 por ciento) y el espesor se había reducido algo. Era evidente que el estrechamiento y adelgazamiento se estaban propagando a lo largo de la banda. Con la siguiente vuelta del tornillo, la banda comenzó a romperse por la zona de estricción; debióse quizás a una pequeña muesca que había dejado con las tijeras en uno de los lados de la banda. (Podría haber arreglado la rotura con un poco de cinta de embalar, pero la dejé como estaba.)

El proceso de estiramiento de la banda puede seguirse con una representación gráfica de la "tensión convencional" en función de la "deformación unitaria" [*véase la figura 4*]. Esta última, que mide el grado de estiramiento, es el cociente entre la variación de longitud de la banda y la longitud inicial de la misma sin estirar. Por ejemplo, cuando se duplica la longitud de la banda, el alargamiento es del 100 por ciento y la deformación unitaria es de 1,0. La tensión convencional es el cociente entre la fuerza con que se tira de la banda y la superficie de la sección transversal tomada en la cara del gato [*figura 6*]. Tal superficie no varió en el curso del experimento, por lo cual los cambios de valor de la tensión convencional reflejan las variaciones en la intensidad de la fuerza aplicada.

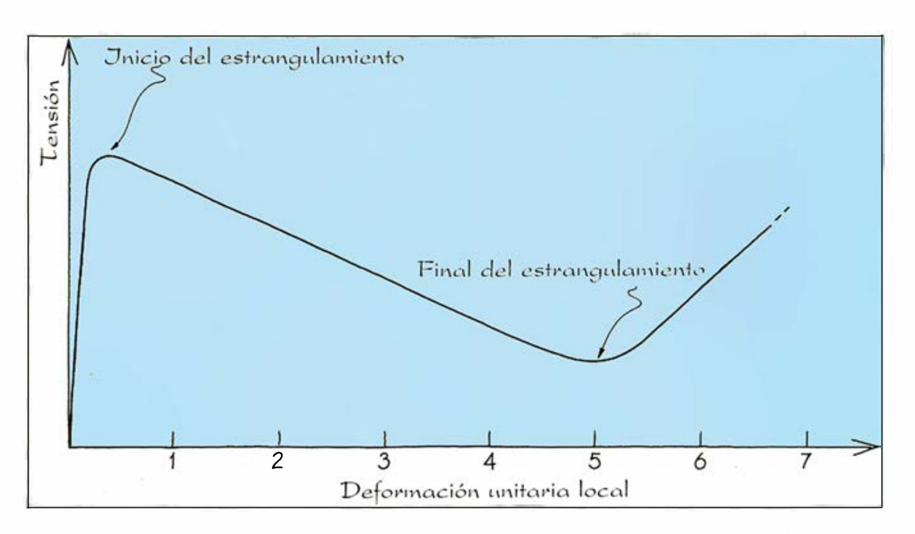



*2. Alineación de las moléculas en la estricción.*

En mi caso, al no poder medir la fuerza de tracción, sólo me fue posible registrar gráficamente mis impresiones subjetivas. Cuando giré el tornillo por primera vez, la tensión se elevó y la banda se estiró marginalmente. Cuando la deformación unitaria llegó a 0,09 comenzó la estricción de la tira, que se alargó bruscamente, con lo que se alivió la tracción




*3. Estiramiento de una tira de polímero.*




*4. Curva de la tensión convencional.*

traban una deformación unitaria local del orden de 4,0.

Regresemos ahora sobre la cuestión de la letra impresa vista a través de una hoja de polímero. El tema fue abordado, en 1973, por David Miller, del Hospital Beth Israel y la Universidad de Harvard, y George B. Benedek, del Instituto de Tecnología de Massachusetts, en su libro *Intraocular Light Scattering*. A este efecto lo llamaron "fenómeno del desnudo en la ducha". (Agradezco la referencia a Craigh F. Bohren, de la Universidad estatal de Pennsylvania.)

Para estudiar la razón del sobrenombre, imaginemos que estamos contemplando la ducha de alguien a través de la cortina de plástico que lo protege (o de una puerta rígida del mismo material). Si la persona se encuentra cercana a la cortina, distinguiremos sin dificultad detalles anatómicos; pero, si se sitúa más lejos, los detalles, muy borrosos, no se distinguirán. Una dependencia similar de la visibilidad con la distancia puede observarse con una cinta adhesiva corriente, de plástico. Sosténgase la cinta (con el lado adhesivo hacia arriba para que no se pegue) encima mismo de esta página, y súbase después gradualmente. Conforme aumente la distancia a la página, las palabras comenzarán a difuminarse hasta tornarse ilegibles.

Para entender cómo se produce la difuminación, centremos la atención en un punto ortográfico de la página. Cuando la luz procedente del punto atraviesa la cinta adhesiva, las moléculas de ésta la difunden [*figura 7, izquierda*]. Merced al proceso de difusión, cada rayo original se abre en un estrecho haz cónico centrado en la dirección del rayo original. Podemos caracterizar cada cono por su semiángulo, es decir, por el ángulo que forman los rayos más desviados del cono con la dirección del rayo original.

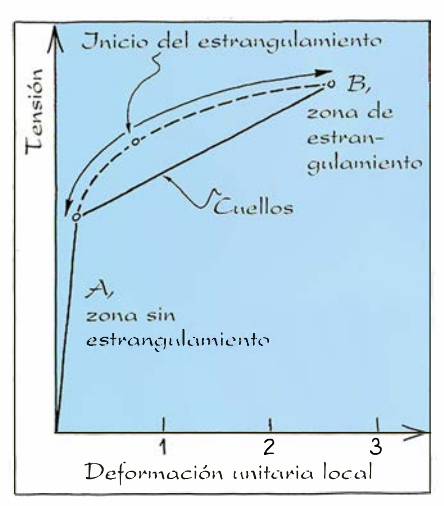
Supongamos que la cinta se encuentre justo encima del punto y que lo miramos con un ojo cerrado, desde una distancia de al menos 40 centímetros. Con el ojo abierto interceptaremos unos rayos que serán aproximadamente perpendiculares a la cinta y que se habrán difundido por las moléculas situadas en la recta que une el ojo con el punto, o por moléculas contiguas a esa recta. Los rayos difundidos por moléculas más apartadas no incidirán en nuestro ojo y no contribuirán a nuestra visión. Al tratar de percibir la fuente de los rayos interceptados, nuestro ojo los extrapolará hacia la página, dando la impresión de que proceden de una mancha centrada en el punto verdadero.

ejercida por el gato y disminuyó la tensión padecida por la banda. Al ir girando el tornillo aún más, aumentando así la deformación unitaria, la estricción se acrecentó y la tensión siguió disminuyendo. Si hubiera sido capaz de estrechar por completo la banda sin que ésta se desgarrara por culpa de las muescas accidentales (muy cerca de ello estuve con varias muestras), la tensión hubiera empezado a aumentar de nuevo conforme las moléculas fuertemente ligadas se opusieran al estiramiento ulterior. Finalmente, la tensión habría partido la banda como si de un hilo quebradizo se tratara.

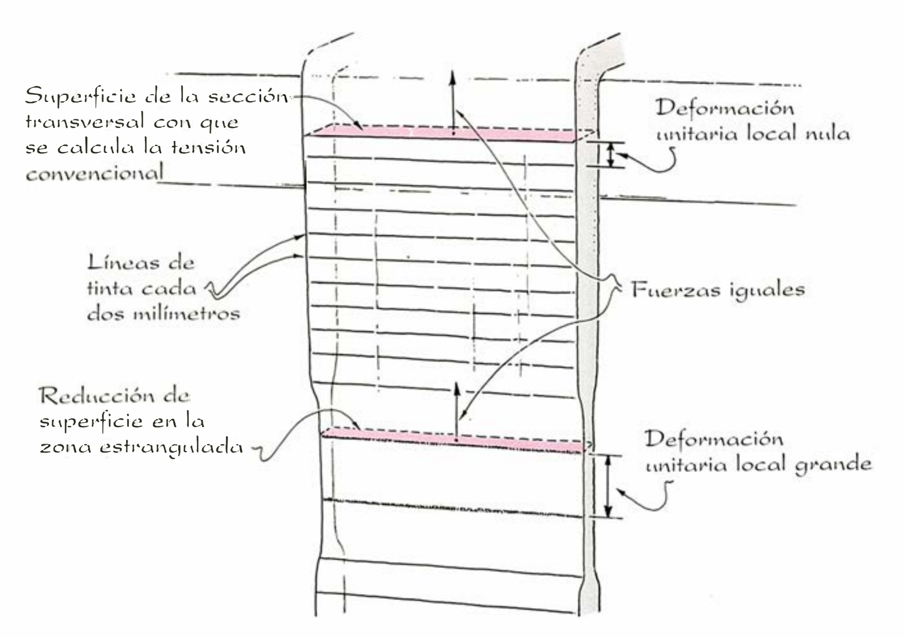
Antes de que aparezca la estricción, la tensión convencional actúa sobre toda la banda, en todas las secciones. Sin embargo, tras la estricción, la tensión en la zona afectada es mayor que la tensión convencional, puesto que en ese lugar la superficie de la sección transversal se reduce. Esa tensión local se llama tensión verdadera y es el cociente entre la fuerza de tracción aplicada en la zona y la superficie de la sección transversal de esa zona. En ésta, la deformación unitaria recibe el nombre de deformación unitaria local.

En la figura 5 se representa la tensión verdadera en función de la deformación unitaria local, antes y después de que aparezca la estricción. Antes de ésta, todas las secciones de la banda sufren la misma tensión y deformación unitaria y el punto representativo de esos valores sube hacia la parte donde se interrumpe la curva. En cuanto aparece la estricción, el punto *A*, representativo de las zonas sin estricción, retrocede curva abajo y el punto *B*, denotativo de la zona afectada, sube aún más curva arriba; o sea, la zona de estricción sufre una tensión y una deformación mayor que las zonas sin estrangular. La línea que enlaza *A* y *B* directamente indica la tensión y la deformación unitaria en los cuellos.

Repetí estos experimentos con bandas de una hoja de polietileno transparente comercializada bajo el nombre de Glad. Lo mismo que el parafilm, las tiras de este envoltorio de alimentos oponían tenaz resistencia al estiramiento inicial cediendo luego bruscamente. Sin embargo, a diferencia del parafilm, presentaban estricción en anchura, de izquierda a derecha, y no en espesor. El estrechamiento, originado en un lugar, se propagaba después gradualmente arriba y abajo de la banda, confiriendo a ésta el aspecto de un deforme reloj de arena. La banda, antes mismo de partirse, presentaba un alargamiento que llegaba al 180 por ciento; su anchura inicial de tres centímetros se había reducido a 0,9 en la zona más estrecha. Ahí, las líneas trazadas a tinta mos-




*5. Curva de la tensión real.*




*6. Factores que determinan la tensión y la deformación unitaria.*

Esa mancha no aparecerá con la misma nitidez que el punto, pero aun así será reconocible. Su radio equivaldrá, aproximadamente, al producto de la distancia punto-cinta por la tangente del semiángulo de los conos de difusión.

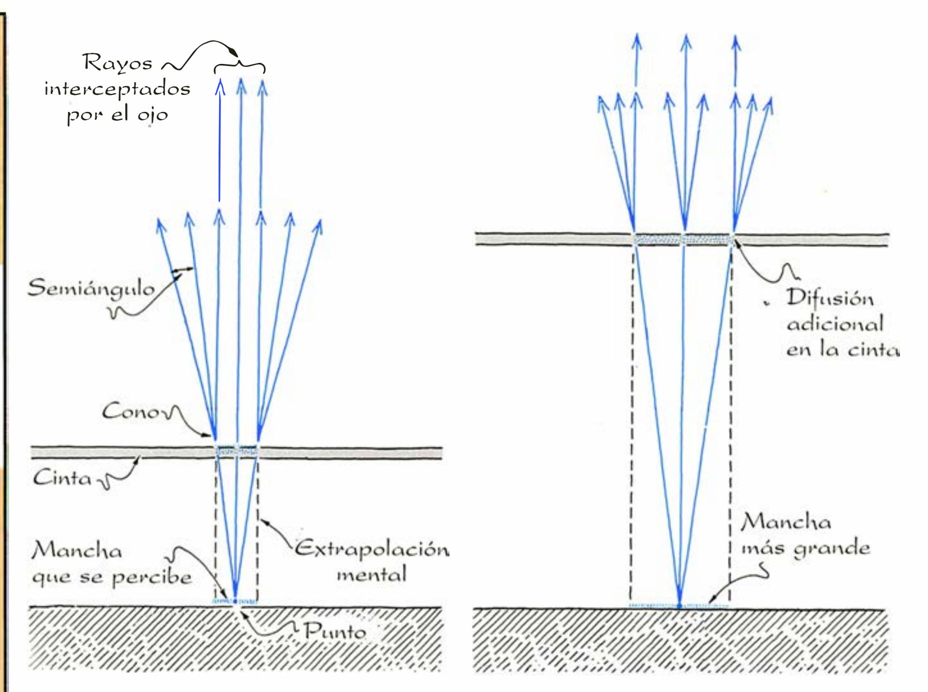
Seguidamente, elevemos la cinta dos o tres milímetros. Tal como se muestra en la ilustración derecha de la figura 7, el semiángulo de cada cono de difusión no varía, pero la mayor distancia punto-cinta esparce más la luz sobre la cinta, con lo que interceptaremos rayos procedentes de una zona de la cinta algo más extensa y también centrada en la recta que une el punto con nuestro ojo. Al extrapolar esos rayos hacia atrás, nos parecerá que salen de una mancha, situada en la página, algo mayor. La mancha será entonces menos clara y resultará más difícil identificarla con el punto que realmente es.

Supongamos que haya un segundo punto junto al primero. Cuando la cinta está próxima a ellos, las manchitas que percibimos son lo bastanque pequeñas y nítidas para que las veamos separadas. Conforme vayamos elevando la cinta y las manchas se vayan ensanchando, acabarán solapándose hasta que resulten indistinguibles. Tal ocurre con el desnudo tras la cortina de plástico: cuando la figura está cerca del mismo, sus detalles se aprecian bien y son reconocibles; pero, cuando la figura está más lejos, los detalles se difuminan.

Al experimentar con parafilm, observé que oscurecía los detalles de una página impresa que se contemplara a su través, igual que la cinta adhesiva. Asumí, por hipótesis, que el grado de difuminación debía depender del espesor de la hoja de plástico. Para cierto espesor de ésta, cada rayo procedente de un detalle dado debe atravesar determinado número de moléculas, sufriendo una difusión en un cono de cierto semiángulo; a este semiángulo corresponde establecer a qué distancia de la página puedo sostener la hoja de plástico sin dejar de percibir el detalle en cuestión. Cabría presumir que, si la hoja adelgaza por la estricción, la luz fuera difundida por menos moléculas y se repartiera así dentro de un cono más estrecho, gracias a lo cual podría yo seguir distinguiendo los detalles con la hoja de plástico más alejada que antes.

Para contrastar esta idea, coloqué un pequeño trozo de parafilm sin estirar sobre uno de los orificios de una plantilla de dibujo técnico y situé el orificio encima de tres pequeños puntos, que marqué con lápiz en un papel. Los puntos formaban una esquina y estaban separados un milímetro vertical y horizontalmente. Icé entonces la plantilla mientras miraba los puntos a través del parafilm. Cuando éste se encontraba a unos dos centímetros del papel, los puntos se difuminaron simultáneamente. Esta observación la repetí varias veces y saqué la media de las "alturas de difuminación".

A continuación sustituí el parafilm por otro trozo que ya había estirado con el gato. Para mi sorpresa, descubrí que los puntos se difuminaban cuando el plástico se encontraba a sólo medio centímetro de ellos. Además, los dos puntos paralelos a la di-




*7. Difusión de la luz en una tira de plástico.*

rección de la estricción se difuminaban alrededor de un milímetro anterior a los dos puntos orientados perpendicularmente a esa dirección. (Desde luego, esa discrepancia podría haberse debido a una separación desigual entre los puntos situados en las direcciones horizontal y vertical, pero persistió incluso cuando giré el papel 90 grados en torno a mi visual.) Aunque mis observaciones fueron sin duda toscas, indican que, cuando el parafilm sufre estricción, difunde la luz en conos cuya anchura viene a cuadruplicar la de los conos que se crean cuando la hoja de plástico no se ha estirado.

En mi opinión, la difusión aumenta a causa de la alineación de las moléculas provocada por la estricción. Cuando la luz se propaga en un medio transparente cuyas moléculas están orientadas al azar, o bien organizadas a una escala mucho menor que la longitud de onda de la luz, se dice que cada molécula difunde la luz independientemente de las otras moléculas. En tal caso, la luz que por el efecto de difusión es desviada con respecto a la dirección de propagación original es probable que sea anulada por la luz difundida en la misma dirección por otra molécula. (Es decir, cada cresta de una onda coincide con un valle de otra onda y una y otra se interfieren destructivamente.) Por tanto, el semiángulo del cono de difusión se mantiene pequeño.

Pero si, por contra, las moléculas están distribuidas siguiendo cierta organización, y si el espaciado asociado a esa organización es del mismo orden que el valor de la longitud de onda de la luz, las moléculas no difunden la luz independientemente, sino de una manera organizada. Cuando dos ondas luminosas son difundidas en la misma dirección, las crestas de una onda quizá no coincidan exactamente con los valles de la segunda, la anulación de la luz difundida hacia ese lado es menos completa y el semiángulo de cada cono de difusión se hace más ancho. Evidentemente, al alinear yo las moléculas provocando estricción en el parafilm, lo que hacía era organizar el fenómeno de difusión, abriendo el cono de difusión y aumentando el emborronamiento producido por la hoja de plástico.


BIBLIOGRAFIA COMPLEMENTARIA

X-RAY DIFFRACTION STUDIES OF THE STRETCHING AND RELAXING OF POLYETHYLENE. Alexander Brown en *Journal of Applied Physics*, vol. 20, n.° 6, págs. 552-558; junio de 1949.

THE NECKING AND COLD-DRAWING OF RIGID PLASTICS. P. I. Vincent en *Polymer*, vol. 1, n.° 1, págs. 7-19; 1960.

ON THE EXTENSION OF THE NECK OF POLYMER SPECIMENS UNDER TENSION. G. I. Barenblatt en *Journal of Applied Mathematics and Mechanics*, vol. 28, n.° 6, págs. 1264-1276; 1964.

INTRAOCULAR LIGHT SCATTERING: THEORY AND CLINICAL APPLICATION. David Miller y George Benedek. Charles C. Thomas, Publisher, 1973.

# Juegos matemáticos

*Las matemáticas permiten vencer sin correr peligro. Tal es su gloria*

Ian Stewart

"Veintisiete", anunció Mabel, la guardiana de las ocas. El dado, al que había hecho rodar un cuarto de vuelta, mostraba un 3.

Abel, el pastor, torció el gesto.

"¡Maldita sea! ¡Ahora no puedo sacar un 4, que me hubiera hecho ganar!"

—Precisamente por eso he jugado así— proclamó Mabel. Quien primero pase de 31, pierde. Como el dado no puede rodar más que un cuarto de vuelta, tus únicas jugadas son ahora 1, 2, 5 o 6. Si juegas al 5 o al 6, pierdes, porque el total excede de 31. Si juegas al 1, el total sube a 28; entonces saco un 3 y gano. Y si juegas al 2, el total pasa a ser 29; yo sacaré entonces un 1, el total será de 30, y, como tú no puedes sacar un 1, yo gano. ¡Así que yo gano! [*Pueden verse las reglas del juego en el recuadro 1.*]

Abel suspiró y se tendió en el suelo.

"Has vuelto a ganarme", reconoció, mordisqueando una brizna de hierba.

—Te lo había advertido ya, exclamó Mabel. Mi lógica es implacable; sabía de antemano que iba a ganar en el momento mismo de empezar. No tenías la más mínima posibilidad. ¿Echamos otra partida?

—No, gracias. ¿Qué tiene que ver la lógica con los juegos? Cuando jugamos al tenis, normalmente soy yo quien gana.

—El tenis es una actividad física, y tú eres más fuerte que yo. Pero en este caso se trata de inteligencia pura: yo he ganado con plena autoridad, merced a mi inexorable estrategia.

—Estrategia... Mucho se oye hablar de esa palabra, pero la verdad es que nunca he entendido del todo a qué se refiere. Supongo que será una forma de llevar bien el juego propio.

—En juegos de esta clase, consiste en ganar haga lo que haga el adversario, dijo Mabel. En ciertos juegos existen estrategias de victoria para el primer jugador; en otros, las hay para el segundo, y en otros más, como el tenis, supongo que carecerán de una estrategia perfecta. La lógica nos señala la estrategia de victoria, *si* es que hay alguna.

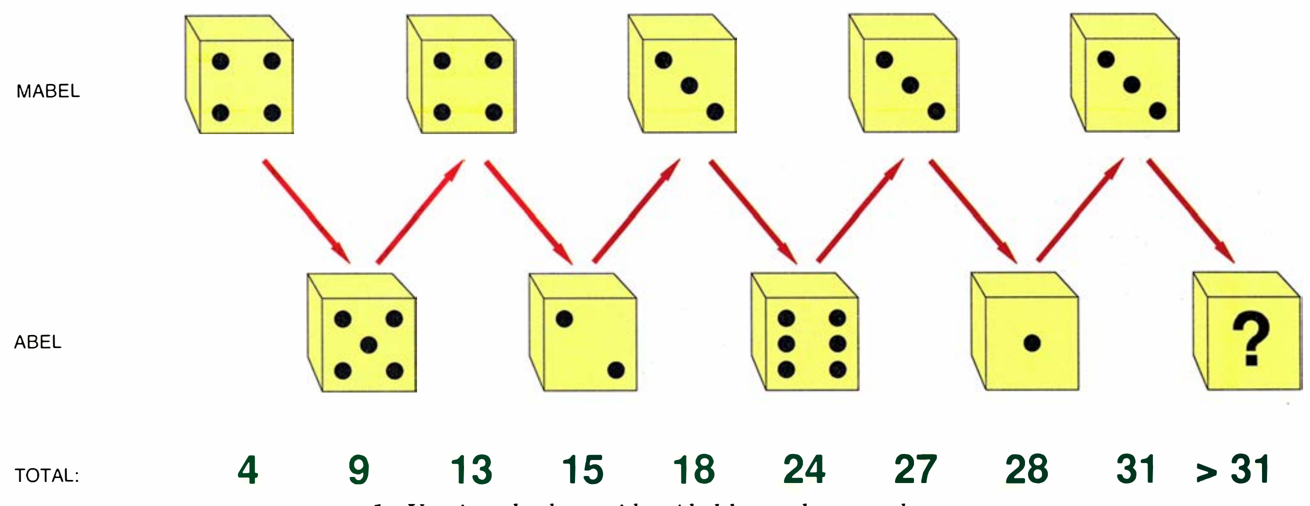
Mabel estiró sobre la hierba sus delgadas piernas bronceadas y se estiró la falda de colores vivos, volviendo a cubrirse los tobillos.

"Examinemos los datos del problema" dijo. "A tal fin, vamos a fijarnos en todos los juegos que verifican cuatro sencillas reglas:

1. El juego tiene tan sólo un número finito de posiciones posibles.
2. Cada partida concluye al cabo de un número finito de jugadas.
3. El juego acaba necesariamente con la victoria de uno de los jugadores. (No hay partidas nulas o "tablas".)
4. Es la posición del juego la que determina las jugadas lícitas subsiguientes, independientemente de cuál sea el jugador de turno."

*El lector puede comprobar que, en el caso de Mabel y Abel, se verifican estas cuatro condiciones. En este juego, las "posiciones" consisten en dos números: el señalado por la cara superior del dado y la suma total de las puntuaciones precedentes. ¿Cuáles de las cuatro reglas anteriores se verifican en el juego de ajedrez?*

"En todo juego de este tipo", prosiguió Mabel, "existe siempre una estrategia que garantiza la victoria bien al primero, bien al segundo de los jugadores. Como mi bondad no desmerece de mi belleza y de mi modes-



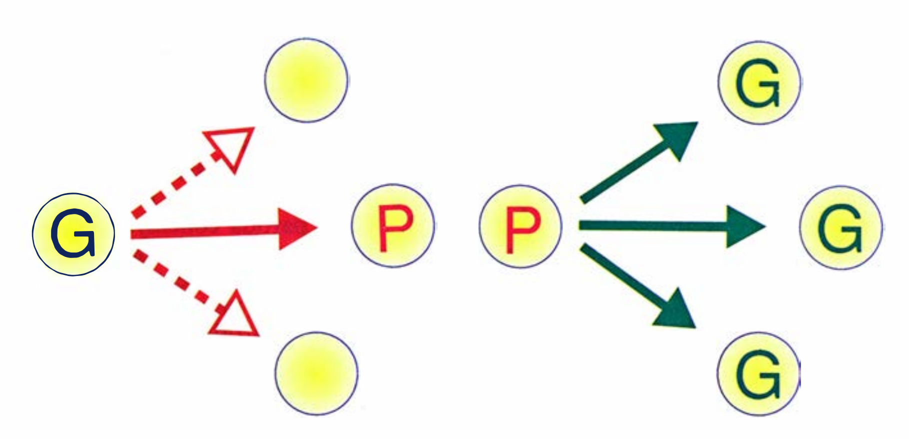

*1. Un ejemplo de partida. Abel ha vuelto a perder.*

## Recuadro 1. Reglas del juego de Mabel y Abel

El primer jugador coloca sobre la mesa un dado ordinario, o sea, un cubo cuyas caras están marcadas de 1 a 6. La puntuación está dada por la cara superior.

Los jugadores actúan por turno, haciendo girar el dado un cuarto de vuelta, lo que lleva a la cara superior una de las cuatro caras verticales. El valor de la cara se suma al total.

El primer jugador que llega a un valor superior a 31 (o a otro total convenido de antemano) pierde la partida.




2. *Posición ganadora* (G) *y posición perdedora* (P).

tia, te voy a enseñar cómo se puede ganar en todo juego que verifique estas cuatro reglas. La idea consiste en retrogradar a partir de la posición ganadora. Si al hacerlo te encuentras en una posición perdedora, entonces tú vas a perder. Con eso quiero decir que el adversario dispone de una estrategia que le dará la victoria. Claro, puede que sea estúpido y no juegue correctamente, pero entre lógicos perfectos...

—Sí, pero... ¿Cómo saberlo?

—Una posición de victoria es una posición que te permite elegir una jugada que situará a tu adversario en una posición perdedora. Por el contrario, las posiciones perdedoras son aquellas tales que, cualquiera que sea la jugada por la que se opte, el adversario queda situado en posición de victoria [*véase la figura 2*]. Ah, se me olvidaba una última cosa: las posiciones finales (regla 3) deben quedar definidas como posiciones perdedoras o ganadoras por las propias reglas del juego. En nuestro juego, por ejemplo, el total de 31 corresponde a una posición ganadora.

—No lo entiendo del todo —la interrumpió Abel—. Para empezar, tú defines "posición ganadora" recurriendo a la idea de "posición perdedora", y luego defines la "posición perdedora" en función de la "posición ganadora". ¡Pero eso es un círculo vicioso!

—¡Nada de eso! —opuso Mabel—. Aunque hay que admitir que a primera vista... En fin, me he valido de una definición por recurrencia. Si sales de una posición cualquiera y sigues una cadena lógica de jugadas, llegarás a fin de cuentas (regla 2) a una posición final de la que nosotros sabemos que es, bien ganadora, bien perdedora. Remontando las series de jugadas, se determinan las posiciones de victoria y derrota.

La chica sacó entonces del bolsillo una tableta de chocolate [*véase la figura 3*].

"La pastilla de chocolate verde ha sido objeto de una preparación que le ha dado un sabor —y aspecto— repulsivo. Las reglas del juego son muy sencillas:

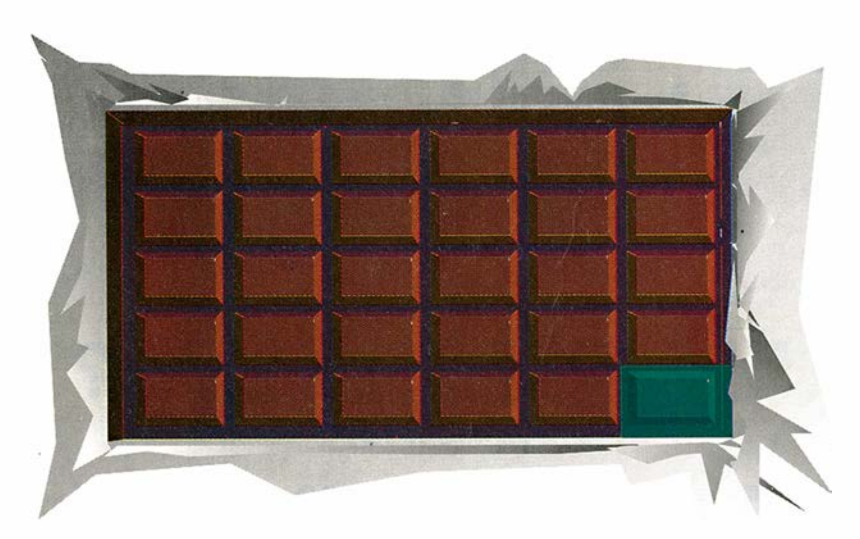
1. Los jugadores van cogiendo la tableta por turno, y separan de ella una parte, cortándola según una sola de las líneas horizontales o verticales que separan las pastillas. Cada jugador se come el pedazo que ha separado.

2. Quien haya de comerse la pastilla verde pierde.

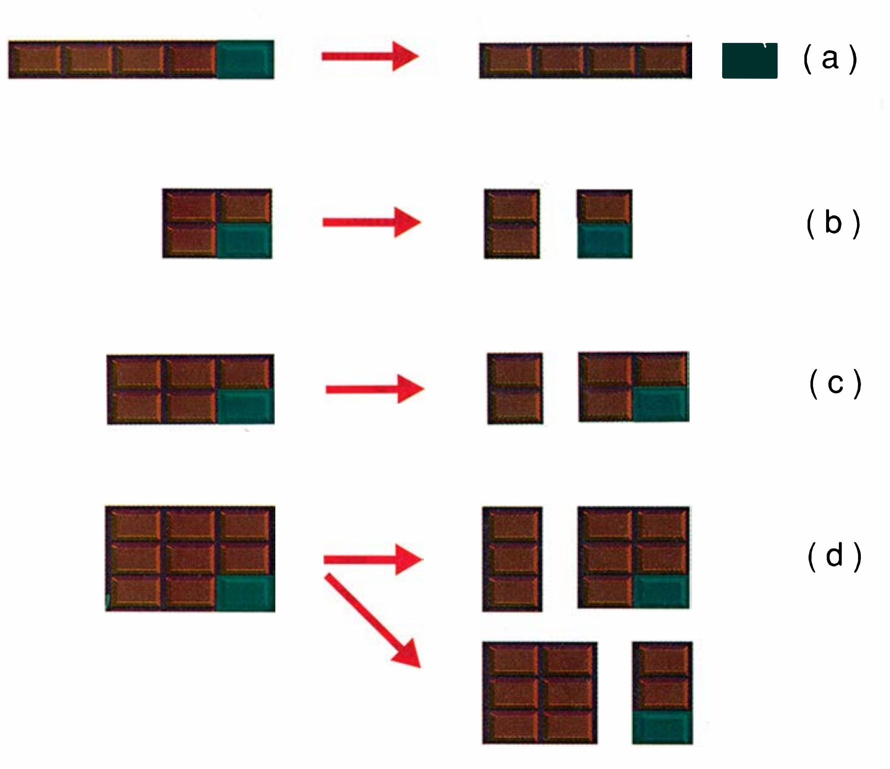
Una posición del juego queda caracterizada, en cada instante, por las dimensiones de la pieza restante de tableta. La posición actual, la de salida, es de 5 × 6. Empieza tú...

—Hum... Esa pastilla verde de los diablos me inquieta...

—Como te decía, opera a la inversa, remontando el tiempo. Si solamente te queda la pastilla verde, te encuentras en una posición perdedora. En consecuencia, toda posición que conduzca a ella en una sola jugada será una posición ganadora. Ahora bien, tales posiciones son precisamente las que dejan una tira de chocolate de grosor 1, o sea, del tipo 1 × 2, o 1 × 3, o 1 × 4, o en general, de 1 × $n$, con $n$ superior a 1. ¿Te das cuenta de esto? Por ejemplo, si recibes una tira de 1 × 5, puedes separar y comerte una banda de 1 × 4 y dejar solamente la pastilla verde [*véase la figura 4a*]. Además, las posiciones 1 × $n$ son las únicas que conducen a 1 × 1 en una sola jugada.



3. *¡No se coma la pastilla de chocolate verde!*




*4. Construcción de una estrategia por retrogradación de la secuencia de posiciones.*

—¿Y las posiciones 2 × 1, 3 × 1,...?

—Tienes razón; pero en este caso no hago distinción alguna entre $a \times b$ y $b \times a$.

—¡Oh! Entonces, si... Espera un momento. Estoy imaginando la poción inicial. No es del tipo 1 × *algo* y, por tanto, yo pierdo, ¿no es verdad? ¡Has vuelto a ganarme, y eso antes mismo de que me tocara a mí!

—Por ahora no sabemos nada. Todavía no nos hemos remontado lo suficiente por la cadena lógica. Veamos el eslabón anterior. Toda posición que conduzca inmediatamente después a una posición de victoria será siempre una posición perdedora. ¿Puedes ver una posición que conduzca siempre a algo dotado de la forma 1 × $n$, con $n$ mayor que 1?

—No, lástima, lástima, lástima.

—2 × 2.

—¿Por qué?

—Porque la única jugada posible es cortarla en dos [*véase la figura 4b*], y dejar dos piezas de 1 × 2.

—¡De acuerdo!

—Yo no veo que haya más posiciones perdedoras en esta fase. Volvamos pues a las posiciones ganadoras. Las posiciones de victoria serán aquellas que lleven a 2 × 2 en una sola jugada. ¿Sabrías describir cuáles son?

—Esto... 2 × 3, porque yo puedo separar 2 × 1 [*véase la figura 4c*], y también 2 × 4, 2 × 5,...

—¡Exactamente! En general, todas las 2 × $n$, con $n$ mayor que 2, y ninguna otra. ¿Y ahora?

—Vamos a buscar una nueva posición perdedora, una posición desde la que siempre se pueda llegar a posiciones que sabemos ganadoras... Veo una situación perdedora: 3 × 3 [*véase la figura 4d*].

—¡Demostrado: Abel es listo! Busquemos ahora nuevas posiciones ganadoras que llevan a 3 × 3, a saber: 3 × 4, 3 × 5,... Y en general, 3 × $n$, con $n$ mayor que 3.

—Observo que hay una regla general...

—Yo también me he dado cuenta.

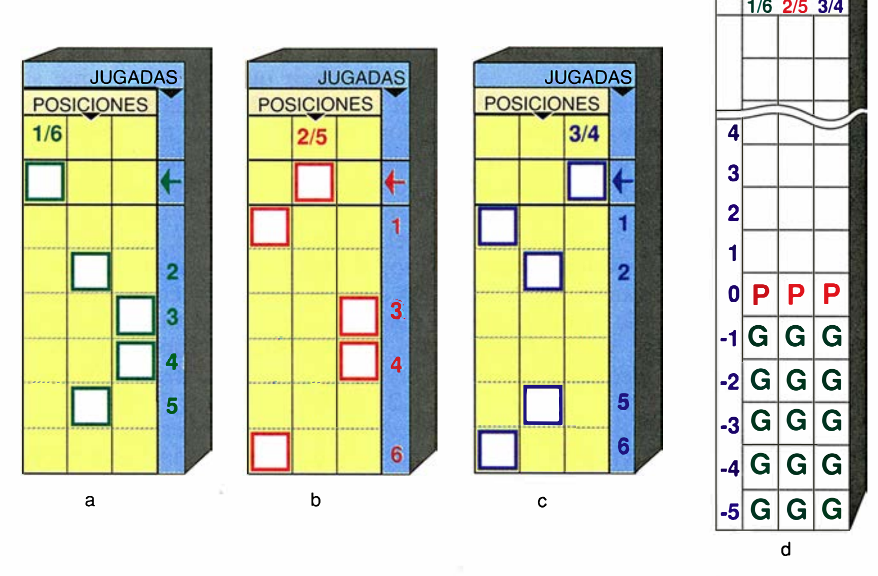
*¿Y usted? ¿En qué consiste esta regla? ¿Puede ganar Abel si comienza con la tableta de 5 × 6?*

Examinemos ahora otro juego del mismo tipo, llamado "Divide y vencerás" [*véase la figura 6*]. Al principio, hay dos cajas que contienen cada una un número no nulo de cuentas. Los jugadores, turno a turno, eliminan el contenido de una de las cajas y reparten el contenido de la otra entre las dos, habiendo cada caja de recibir una cuenta cuando menos. El juego se detiene cuando uno de los jugadores no puede actuar lícitamente, de acuerdo con las reglas, y, por tanto, pierde. ¿Cuáles son las posiciones ganadoras y cuáles las perdedoras en este juego?

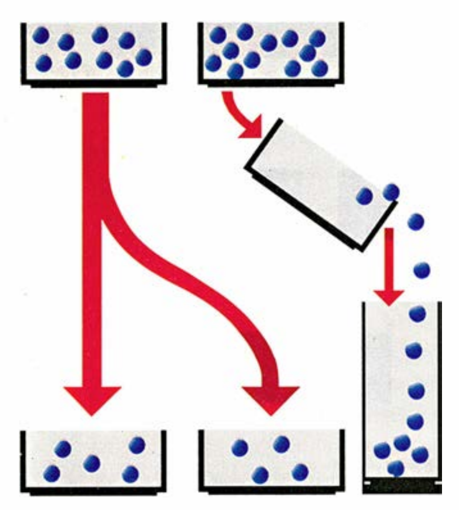
"El juego de voltear el dado se puede estudiar de igual manera que el de la tableta de chocolate, prosiguió Mabel. El total de 31 elegido por nosotros nada tiene de particular y podemos estudiar una estrategia de victoria, cualquiera que sea el total convenido de antemano.

Como has visto, cada posición del juego puede ser descrita mediante dos números: la puntuación de la cara superior y el total actual. Por mor de sencillez, voy a sustituirlos por otros dos.

Para empezar, en lugar del total actual utilizaré la diferencia entre ese total y el límite convenido; dicho de




5. (a-c) *Las fichas estratégicas: recorte las ventanas correspondientes a los cuadrados blancos.* (d) *Así se comienza a rellenar el casillero.*



*6. Un ejemplo de jugada en el juego de "Divide y vencerás".*

otra forma, el número de puntos que faltan antes de que uno de los jugadores rebase ese límite. Llamaré a ese número "suma pendiente"; gracias a ella resultará fácil remontar las cadenas lógicas.

Observo enseguida que las jugadas posibles, cuando la cara 1 está en lo alto (a saber, las que llevan a 2, 3, 4 o 5) son las mismas que cuando la cara visible marca 6 (pues el total de puntos de cada par de caras opuestas es 7). Los movimientos que parten de caras opuestas forman una "cinta" sobre el dado, de suerte que las posiciones "cara 1 en lo alto" y "cara 6 en lo alto" pueden considerarse idénticas. Otro tanto vale para las caras 2 y 5; también para las 3 y 4. Por consiguiente, no distinguiré más que tres posiciones: 1/6, 2/5, o 3/4.

—¡Un momento! Si llevo a lo alto la cara 6, el total aumenta en 6 unidades; en cambio, si es la cara 1, solamente aumenta en 1.

—Sí, pero en ese caso estás refiriéndote a jugadas, no a posiciones. Estoy de acuerdo contigo en que al hablar de jugadas hay que distinguir las caras opuestas.

—Desde luego.

—El proceso regresivo para descubrir las posiciones ganadoras o perdedoras es complicado y resulta fácil equivocarse. Por eso voy a fabricar ahora mismo una pequeña calculadora analógica."

La chica sacó unas tijeras y recortó tres fichas como las representadas en las figuras 5*a* a 5*c*. Después, dibujó tres columnas en una hoja de papel, como las que vemos en la figura 5*d*.

"El casillero representa las posiciones posibles del juego. Cada una de las tres columnas corresponde a dos posiciones equivalentes del dado; cada línea está asociada con el total restante. He llenado la línea cero con letras *P*; en efecto, si el total que falta es cero, las tres posiciones son perdedoras para el jugador que esté de turno.

—¿Qué significado tienen las *G* que están debajo?

—No es más que un convenio, que facilita el comienzo del cálculo. Convendrás en que, si heredas un total negativo, forzosamente tendrás que haber ganado, pues la jugada precedente ha hecho perder a tu adversario. También los totales negativos han de ser considerados posiciones de victoria.

—Admitámoslo...

—Fíjate bien en la secuencia de operaciones y pronto verás que es razonable —afirma Mabel—. Todavía no te he explicado el papel de las fichas. Hay una para cada una de las posibles caras superiores: 1/6, 2/5 o 3/4. Tomemos la pertinente a 1/6. La perforación situada debajo mismo de 1/6 te permite escribir en el casillero. Las otras cuatro perforaciones indican la forma en que las demás jugadas posibles (2, 3, 4 o 5) modifican el total que falta. Si juegas 2, el total que falta se reduce en 2 unidades, y la perforación correspondiente se encuentra dos líneas más abajo, y así en los demás casos.

—Me parece que ya veo.

—Puedo utilizar las fichas para "retrogradar" en el juego. En este momento, para hallar las jugadas de victoria sobre una línea dada coloco sucesivamente las tres fichas sobre el casillero, de modo que los orificios de escritura, los indicados por las flechas, se encuentren sobre esta línea. Si para una ficha dada no aparece una *P* en los otros cuatro orificios, eso quiere decir que cualquiera que sea mi jugada, ésta conduce a una posición victoriosa para mi adversario. Esa es precisamente la definición de posición perdedora, así que yo marcaré sobre el casillero una *P* a través del hueco de escritura. Por el contrario, si aparece cuando menos una *P*, es que me encuentro en una posición ganadora para mí, y esta es la jugada (o jugadas) que yo anotaré en la casilla. Fíjate en cómo lleno las líneas 1 y 2 [*véanse las figuras 7a y 7b*].

Al proseguir con esta operación, nos vamos remontando hacia totales restantes cada vez mayores [*véase la figura 8*]. Hemos encontrado así una estrategia perfecta: para cierta posición, es decir, dada una fila y una columna, el casillero indicará *P* o no. Si señala una *P*, habrás perdido frente a un jugador que utilice el mismo casillero y desarrolle un juego perfecto. Por el contrario, si no es una *P*, puedes realizar cualquiera de las jugadas indicadas, a sabiendas de que conducen a posiciones perdedoras para tu adversario.

Queda una última cuestión que conviene no olvidar. La primera posición y la primera jugada son, en parte, arbitrarias. Por ejemplo, si soy yo quien empieza y el total restante es de 31, puedo crear las posiciones representadas en la tabla correspondiente. Tú puedes utilizar el casillero para saber cuáles son ganadoras y, cuáles, perdedoras.

Puedes elegir la cara que se encuentra arriba en la primera jugada, pues no está determinada por ninguna posición antecedente. De hecho, la línea 31 es 4 4 *P*. La posición ganadora es 4. Quien haya de jugar a partir de la posición 4 (y de total 4) habrá perdido si su oponente juega correctamente en lo sucesivo.

Cuando nosotros jugamos [*véase la figura 1*], el total restante en la primera jugada era 31 y fui yo la primera en actuar. Como es obvio, yo sabía cuanto te acabo de explicar; consiguientemente, sabía que tenía que jugar 4. Al hacerlo, el resto total quedaba rebajado a 27. Entonces tú jugaste 5, dejándolo en 22. Según el casillero, la posición línea 22, columna 2/5 es ganadora, y la jugada a efectuar es 4. ¡Que fue lo que hice, claro está! Tú respondiste con un 2, y yo me encontré en la fila 16, columna 2/5, que me daba a elegir entre 3 y 4; opté por el 3. Tú jugaste 6, lo que me llevó a la línea 7, columna 1/6. En ella, 2, 3 y 4 son ganadores; volví a escoger el 3. Eso te impedía elegir el 4 (y ganar), así que escogiste el 1. Entonces yo hice mi jugada final, otro 3, y quedaste liquidado.

—¡Muy inteligente! ¡Pero si el resto total fuese más grande te haría falta una lista kilométrica!

—No es así. ¿No has observado nada interesante en esta lista?

—No."

*¿Y usted, lector? Reflexione antes de proseguir.*

Los resultados a partir de 17 no hacen más que repetir lo que podemos leer a partir de 8. Dicho de otro modo, el casillero es periódico, con período 9. A partir de la octava, cada línea se repite 9 líneas más tarde. Así, la estrategia para un resto total de 31 no se diferencia de la correspondiente a 31 − 9 = 22, o de la de 22 − 9 = 13, y así sucesivamente. Y la correspondiente a 1012 es la misma

que la de 1012 − (9 × 111) = 13. ¡Es suficiente recordar la lista de las 17 líneas finales del casillero!

"Aún así es un latazo, suspiró Abel, para quien aprender de memoria una sola línea suponía un esfuerzo heroico.

—Un matemático diría que la estructura está determinada por el total que falta módulo 9 —añadió Mabel—, es decir, por su resto en la división por 9, o sea, prácticamente, por el intervalo 8-16, pues las primeras líneas rompen la estructura, como tantas veces ocurre. En resumen, tienes que operar así:

Si el total inicial (el resto) es 8 o menos que 8, lo conservas.

En caso contrario, formas su "raíz numérica" sumando sus cifras reiteradamente hasta conseguir un número comprendido entre 1 y 9 (inclusive). Obtienes así el resto de la división por 9; si el resto fuese 0, lo sustituyes por 9.

Si la "raíz" es 8 o 9, la conservas.

De no ser así, le sumas 9.

Resulta así que la estrategia para un resto de 10 es distinta de la correspondiente a un 1, mientras que para 19 es la misma que para 10. Este resultado determina la fila de que nos valdremos para aplicar la estrategia.

Partamos, por ejemplo, de un resto total de 1012. Es mayor que 8, no cabe duda. Su "raíz digital" es 1 + 0 + 1 + 2 = 4. No es ni 8 ni 9, y, por tanto, le sumamos 9, lo que da 13. Así, pues, una estrategia ganadora para un resto total de 1012 está dada por la línea 13 del casillero, como ya te había dicho.

—¡Soberbio!

—¡Por fin te das cuenta de la potencia de los métodos recursivos! Gracias a ellos, hemos resuelto por completo el problema, cualquiera que sea el "resto total inicial". Nuestra estrategia perfecta resulta bastante más difícil de comprender en su conjunto que el problema de la tableta de chocolate. No creo que hubieras podido descubrirla tú solo.

—No tienes mucha confianza en mi capacidad, ¿no es así, Mabel?

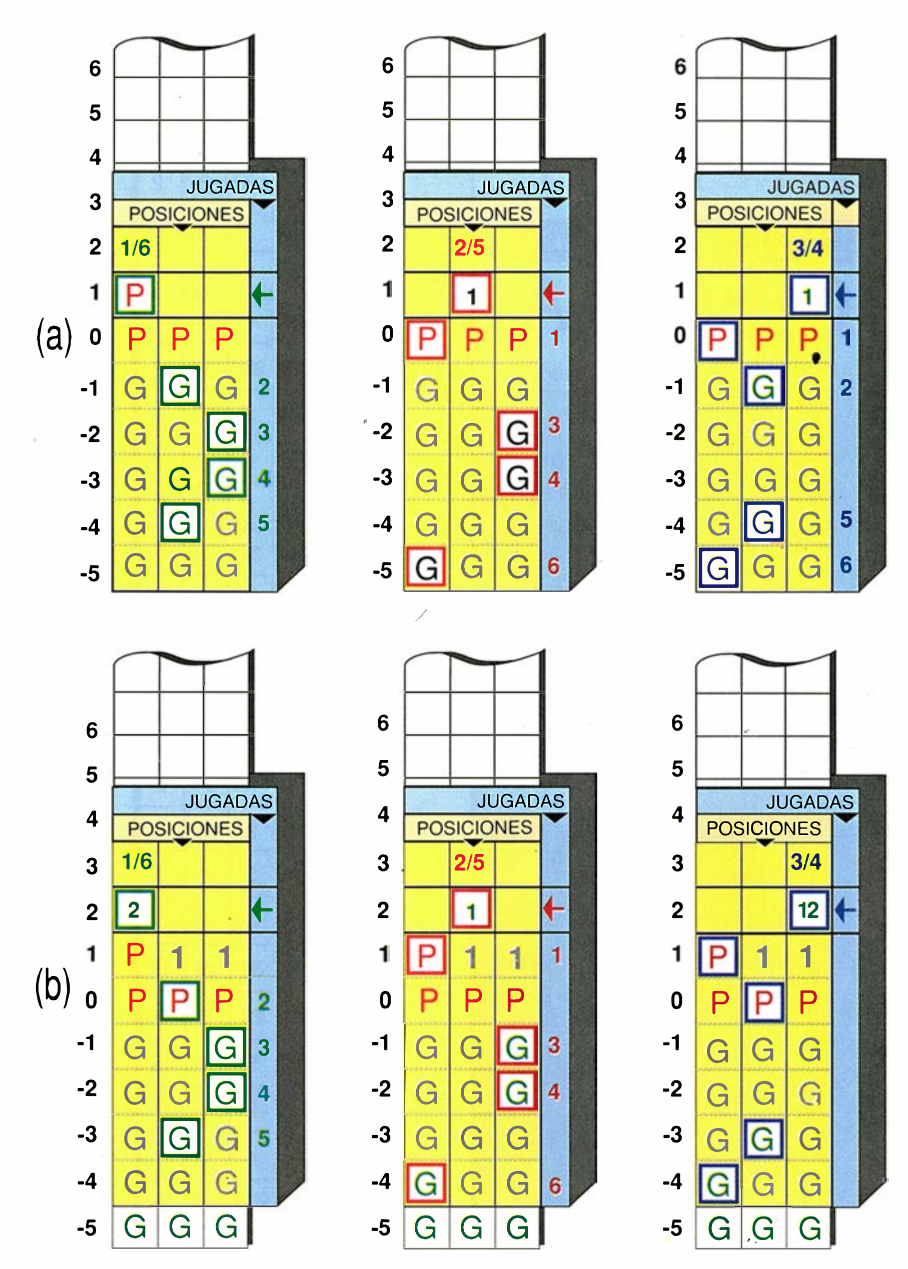
—¡Es que tampoco la he inventado yo, Abel! Pero ahora, para comprobar hasta qué punto la has entendido, voy a pedirte que idees un conjunto de fichas apropiadas y que prepares una estrategia dominante para ese mismo juego, suponiendo que el dado sea tetraédrico, con las caras numeradas de 1 a 4. Esta vez será la cara inferior la que sumaremos al total, pues no hay cara superior."

Abel, superó tan dura prueba. *¿Podría usted hacer otro tanto*? Le daré una pista: cuando el dado es tetraédrico, la estructura es periódica, de período 10. *En el caso general de los juegos de este tipo, en el que se va haciendo rodar cierto objeto de caras numeradas con el fin de modificar un total, ¿será siempre periódico el casillero de estrategias*? De ser la respuesta afirmativa, existirá una estrategia válida cualquiera que sea el total inicial, descriptible en una cantidad limitada de papel.

## Soluciones

*¿Se respetan las reglas?*

*(a) En el juego de Mabel y Abel.* Con un límite de 31 puntos y seis caras en el dado, el número total de po-




*7.* (*a*) *Para determinar la estrategia a utilizar en la línea 1, se coloca la ficha correspondiente a la posición 1/6 con la flecha sobre la línea 1 y se examinan las cuatro ventanas que hay debajo. Todas ellas indican* G, *con lo que la posición considerada es perdedora: por tanto, se escribe* P *en la ventana indicada por la flecha. Procédase de igual manera con la ficha 2/5. Esta vez, aparece una* P *en la ventana correspondiente a "jugar 1", con lo cual "jugar 1" es una jugada que da la victoria: se escribe un 1 en la ventana señalada. Se repite la misma operación con la ficha 3/4: lo mismo que antes, 1 es una jugada ganadora.* (b) *Se pasa ahora a la línea 2, y así sucesivamente. Se puede ascender por el casillero tanto como se desee.*

| TOTAL RESTANTE | 1/6 | 2/5 | 3/4 |
|---|---|---|---|
| 31 | 4 | 4 | P |
| 30 | 34 | 34 | P |
| 29 | 23 | 3 | 2 |
| 28 | 5 | 1 | 15 |
| 27 | P | P | P |
| 26 | 4 | 4 | P |
| 25 | 234 | 34 | 2 |
| 24 | 3 | 36 | 6 |
| 23 | 5 | P | 5 |
| 22 | 4 | 4 | P |
| 21 | 34 | 34 | P |
| 20 | 23 | 3 | 2 |
| 19 | 5 | 1 | 15 |
| 18 | P | P | P |
| 17 | 4 | 4 | P |
| 16 | 234 | 34 | 2 |
| 15 | 3 | 36 | 6 |
| 14 | 5 | P | 5 |
| 13 | 4 | 4 | P |
| 12 | 34 | 34 | P |
| 11 | 23 | 3 | 2 |
| 10 | 5 | 1 | 15 |
| 9 | P | P | P |
| 8 | 4 | 4 | P |
| 7 | 234 | 346 | 26 |
| 6 | 3 | 36 | 6 |
| 5 | 5 | P | 5 |
| 4 | 4 | 4 | P |
| 3 | 3 | 3 | P |
| 2 | 2 | 1 | 12 |
| 1 | P | 1 | 1 |
| 0 | P | P | P |
| -1 | G | G | G |
| -2 | G | G | G |
| -3 | G | G | G |
| -4 | G | G | G |
| -5 | G | G | G |

*8. Casillero estratégico.*

siciones posibles es, a lo sumo, de $32 \times 6 = 192$ (32, porque 0 es el total de salida y 31 es una puntuación aceptable). Se verifica, pues, la regla 1. Como el total aumenta al menos en una unidad en cada jugada, el juego no podrá durar más de 32 jugadas; de hecho, serán menos, ya que no se puede jugar el 1 dos veces seguidas. La sucesión de jugadas 1, 2, 1, 2, ..., 1, 2, 1, que consta de 21 términos, es la partida más larga aceptable. Sea como fuere, se respeta la regla 2; es obvio que otro tanto vale para las reglas 3 y 4.

*(b) Al ajedrez*. Puede haber partidas nulas (tablas); por tanto, no se respeta la regla 3. Tampoco se respeta la regla 4 (y por eso los enunciados de los problemas de ajedrez especifican siempre si les corresponde jugar a las negras o a las blancas). No obstante, la regla 4 sí se verifica definiendo nuevamente las "posiciones" del juego como pares ($P$, $Q$), donde $P$ representa la disposición de las piezas sobre el tablero y $Q$ especifica quién debe efectuar la jugada siguiente. La regla 1 sí se respeta: con 32 piezas y 64 casillas en el tablero, existen a lo sumo $32^{65}$ posiciones (65, porque una pieza puede quedar eliminada y ser retirada del tablero), número que es aproximadamente igual $6835 \times 10^{97}$. Puede que no se respete la regla 2: en principio, una partida podría continuar indefinidamente (basta desplazar alternativamente cada uno de los reyes de una a otra de dos casillas permitidas); no obstante, una de las reglas del ajedrez estipula entonces que cada jugador puede declarar tablas cuando la misma posición se repite tres veces. La regla no dice que al menos uno de ellos tenga forzosamente que hacerlo, pero formulando la hipótesis de que una partida nula sea siempre declarada como tal, en aplicación de esta regla, el juego concluye en un número finito de jugadas.

### *El chocolate verde*

Las posiciones perdedoras son las formas "cuadradas" $1 \times 1$, $2 \times 2$, ..., $n \times n$. Todas las demás posiciones, las formas "rectangulares", son ganadoras: en una sola jugada podemos transformar todo rectángulo en un cuadrado (toda posición ganadora conduce, si se juega bien, a una posición perdedora), mientras que todo corte efectuado sobre un cuadrado deja un rectángulo (una posición perdedora no puede conducir más que a posiciones ganadoras).

Así pues, Abel parte de una posición de victoria. Su primera jugada ha de consistir en separar la barra de

| FORMA DE COMENZAR | | | |
|---|---|---|---|
| PRIMERA JUGADA | CARA SUPERIOR | TOTAL RESTANTE | POSICION |
| 1 | 1/6 | 30 | *G* (jugar 3 o 4) |
| 2 | 2/5 | 29 | *G* (jugar 3) |
| 3 | 3/4 | 28 | *G* (jugar 1 o 5) |
| 4 | 3/4 | 27 | *P* |
| 5 | 2/5 | 26 | *G* (jugar 4) |
| 6 | 1/6 | 25 | *G* (jugar, 2, 3 o 4) |

$1 \times 5$ y dejar el cuadrado de $5 \times 5$. A partir de ese momento, elija Mabel lo que elija, Abel puede dejar nuevamente una tableta cuadrada (más pequeña), obligando a Mabel a concluir el juego con la "pastilla" verde.

"*Divide y vencerás*"

Debo este juego a Keith Austin.

Las posiciones están dadas por parejas $(m,n)$ de enteros no nulos, correspondientes a $m$ cuentas en una caja y a $n$ cuentas en la otra. Las posiciones de victoria son aquellas en las que al menos uno de los números $m,n$ es par. Las posiciones perdedoras son aquellas en las que $m$ y $n$ son ambos impares.

¡Verifiquémoslo! Partamos de una posición ganadora, en la que al menos uno de los números será par, y podremos jugar así: repartir este número par de cuentas entre las dos cajas (después de haber vaciado la otra), de suerte que cada una contenga un número impar de cuentas (poniendo, por ejemplo, una sola en una de las cajas). Obtendremos de ese modo una posición perdedora (para el adversario), compuesta por dos números impares.

Partamos ahora de una tal posición; cualquiera que sea la caja elegida, el número de cuentas a repartir será impar. Como en todo reparto en dos de un número impar, una de las partes será par y la otra impar (porque impar + impar, lo mismo que par + par, da par). Llegamos, pues, en todos los casos a una posición de victoria.

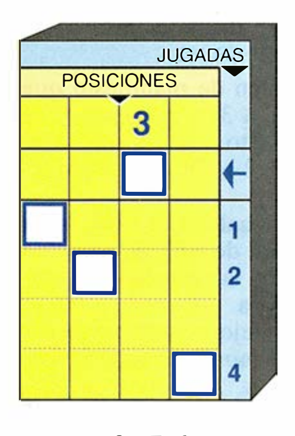
*El dado tetraédrico*

En la figura 9 se representa un conjunto de fichas, y la estrategia resultante, en la figura 10. Observemos





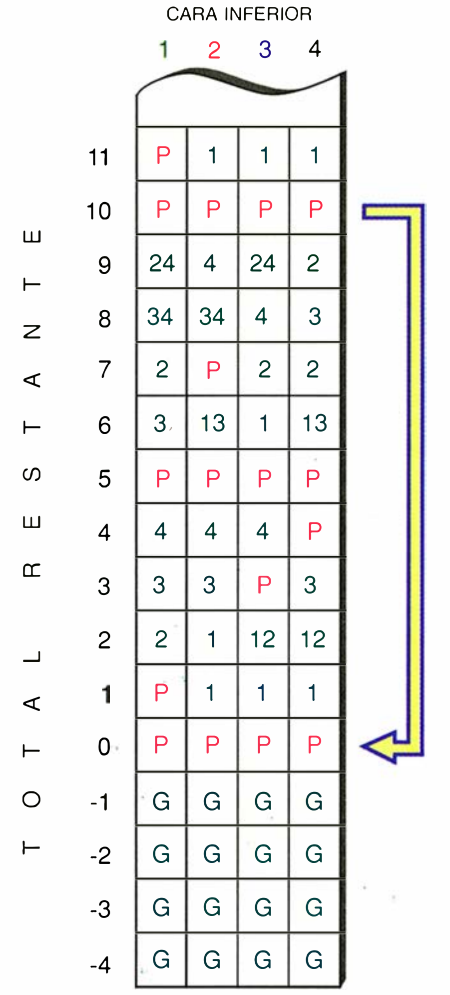






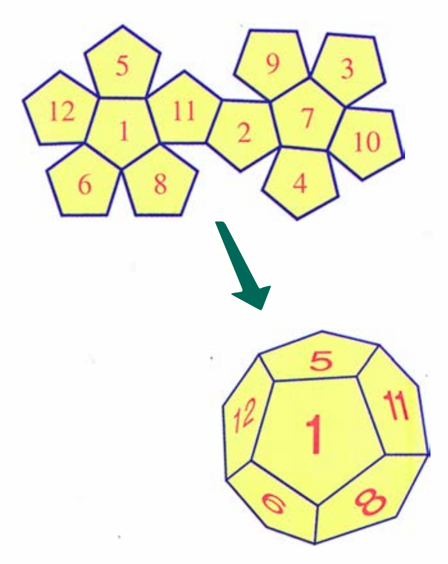


9. *Fichas estratégicas para un dado tetraédrico.*



| TOTAL RESTANTE \ CARA INFERIOR | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 11 | P | 1 | 1 | 1 |
| 10 | P | P | P | P |
| 9 | 24 | 4 | 24 | 2 |
| 8 | 34 | 34 | 4 | 3 |
| 7 | 2 | P | 2 | 2 |
| 6 | 3 | 13 | 1 | 13 |
| 5 | P | P | P | P |
| 4 | 4 | 4 | 4 | P |
| 3 | 3 | 3 | P | 3 |
| 2 | 2 | 1 | 12 | 12 |
| 1 | P | 1 | 1 | 1 |
| 0 | P | P | P | P |
| -1 | G | G | G | G |
| -2 | G | G | G | G |
| -3 | G | G | G | G |
| -4 | G | G | G | G |

*10. El casillero correspondiente a un dado tetraédrico se repite cada diez líneas (el período es 10).*



*11. Un dado dodecaédrico. El casillero correspondiente a 12 columnas tiene que ser periódico. A pesar de ello, ¡no se observa ninguna repetición en las 1000 primeras líneas!*

que la estructura es periódica a partir de la línea 10, como la flecha indica.

*¿Juegos periódicos?*

Para este tipo de juegos, consistentes en numerar las caras de un poliedro y hacerlo rodar al objeto de obtener, por adición, un total prefijado, la respuesta es afirmativa. Supongamos que el mayor de los números inscritos sobre las caras sea $n$. Si un conjunto de $n$ líneas consecutivas de la tabla de la estrategia se repite una vez, la tabla entera será periódica (la ficha de la estrategia no tiene más que $n$ líneas, y en ella no se toman en consideración las secuencias de más de $n$ jugadas). El número de tales conjuntos distintos, aunque posiblemente muy grande, es finito y, por consiguiente, es inevitablemente una repetición.

Poseo un dado dodecaédrico, que utilizo para jugar a "Castillos y dragones" [*véase la figura 11*]. Yo no propondría que se utilizasen fichas de estrategia para estudiar un juego así, sino que se programase un ordenador personal. En virtud del resultado general, el casillero de estretagia ha de ser periódico; a pesar de lo cual, no he descubierto ninguna repetición al escribir explícitamente sus 1000 primeras líneas. ¡Sea cual fuere la estrategia, exigirá un esfuerzo de memorización muy considerable!

*Puesta al día*

Se han obtenido nuevas informaciones concernientes a la función $s(n)$, que es el número más pequeño tal que existan $s(n)$ potencias $n$-ésimas cuya suma sea una potencia $n$-ésima. X. Gonze, de Lovaina, Bélgica, me ha enviado las nuevas igualdades siguientes:

$$12^8 = 2 \times 11^8 + 3 \times 5^8 + 4^8 + 4 \times 3^8 + 2^8 + 23 \times 1^8$$
$$5^9 = 7 \times 4^9 + 6 \times 3^9 + 19 \times 1^9$$
$$7^{11} = 5 \times 6^{11} + 3 \times 5^{11} + 4 \times 4^{11} + 39 \times 2^{11},$$

que ponen de manifiesto que $s(8) \leqslant 34$, $s(9) \leqslant 32$ y $s(11) \leqslant 51$. Gonze ha descubierto también que: $125^6 = 118^6 + 93^6 + 2 \times 78^6 + 48^6 + 42^6 + 18^6 + 2 \times 6^6$.

Esta igualdad, aunque no mejora la acotación de $s(6)$ ya conocida, es notable por la pequeñez de los enteros que utiliza.

Las mejores acotaciones de $s(n)$, hasta el momento actual, son:

| $n$ | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $s(n) <$ | 3 | 4 | 7 | 8 | 34 | 32 | 102 | 51 | 135 | 150 | 166 | 181 |

# Libros

## *Estado metálico, técnica castellano-leonesa, Buffon e interpretaciones del mundo cuántico*

León Garzón, Pedro Molera, Luis Alonso y José L. Sánchez

**ELECTRONS IN METALS AND ALLOYS**, por J. A. Alonso y N. H. March. Academic Press, 1989.

En sentido estricto una teoría del estado metálico permitiría calcular los valores de las diferentes propiedades de los metales y aleaciones a partir de las de sus átomos constituyentes, de alguna magnitud susceptible de determinación experimental o de ambos. Para su establecimiento sería necesario conocer la expresión de las fuerzas que operan en el sistema; y, aunque la naturaleza de estas fuerzas –incluidos los efectos cuánticos– sea bien conocida, sin embargo, el hecho de que dicho sistema consista en una colectividad enorme –del orden del número de Avogadro– de iones y electrones complica extraordinariamente el problema. Ante esta dificultad de principio, es preciso conformarse con modelos de diverso alcance y complejidad, los cuales no son contradictorios sino complementarios. Por otra parte se han desarrollado modelos semiempíricos, que suelen ser muy útiles para el establecimiento de correlaciones entre las propiedades y algún parámetro fundamental.

En su libro, Alonso y March aplican los diferentes modelos existentes con objeto de obtener, en cada caso, descripciones, lo más realistas posibles, del comportamiento de los metales puros, tanto cristalinos como desordenados, así como de las aleaciones.

Consta de 10 capítulos y un apéndice. El primero de ellos es una breve introducción general. En el segundo se expone el modelo de Fermi-Thomas, en el que se llega a establecer la ecuación de estado. En el tercero se presenta el modelo semiempírico de Miedema que, partiendo del concepto de celdilla atómica y de otras hipótesis plausibles, permite establecer correlaciones entre propiedades globales de las aleaciones binarias y las densidades electrónicas en las fronteras de las celdillas de cada constituyente.

Mención especial merece el capítulo 4, en el que, con los conocimientos desarrollados en los capítulos precedentes, se aborda el estudio de las propiedades globales de los metales utilizando tratamientos diversos (pseudo potenciales y teoría de las perturbaciones de primero y segundo orden fundamentalmente). Llama la atención (aunque sobre ello no se detengan los autores) la gran diferencia existente entre la conducta de los metales *sp* y la de las series de transición cortas, es decir, las 3*d*, 4*d* y 5*d*. Me refiero al hecho de que, para las dos últimas series, se advierte claramente las regularidades existentes de acuerdo con la periodicidad, siendo, en general, aceptable, la concordancia obtenida entre los valores calculados (con el modelo) y los experimentales. En cambio, la serie 3*d* es la menos regular, lo que los autores atribuyen a efectos magnéticos de algunos elementos de ella, sin entrar en más detalles. Hubiera sido interesante incluir en este capítulo (aparte de los lantánidos) a los metales del 7.º período, puesto que algunos de éstos, tales como el U, Pu y Am, además de ofrecer un gran interés tecnológico, presentan características muy notables.

En el capítulo 5 se estudian los defectos puntuales, concediéndose suficiente atención a las técnicas experimentales basadas en la reacción de aniquilación entre positrones y ciertos tipos de defectos. El capítulo 6 se dedica a las aleaciones cristalinas, cubriéndose los mismos temas que en el capítulo 3, si bien en este caso se parte de un punto de vista más fundamental, tratándolo de conectar con el modelo de Meidema. Su extensión –un 15 por ciento del total del libro– da una idea de la importancia que los autores han dedicado a este aspecto del estado metálico, lo que sin duda constituye un acierto. En el capítulo 7 se estudian los sistemas metálicos desordenados, de creciente importancia tecnológica.

Finalmente, y como aplicación de los modelos, en los capítulos 8, 9 y 10 se tratan sucesivamente los fenómenos de transporte, del magnetismo y los de superficie. En el apéndice se incluyen cuestiones teóricas relevantes que complementan ciertos desarrollos dados en el texto. El número de referencias bibliográficas es tan enorme que en este sentido pueden ser de gran utilidad para cualquier persona que desee iniciar un trabajo de investigación sobre alguno de los temas tratados. Por consiguiente, estamos ante un libro que abarca prácticamente todos los aspectos de esta rama de la física del estado sólido.

La exposición es suficientemente rigurosa, existiendo, además, un equilibrio entre los niveles teórico y fenoménico, lo que a nuestro juicio constituye un acierto.

En resumen, el libro de Alonso y March puede ser muy útil tanto para científicos como para técnicos que posean vocación hacia estos temas de la materia condensada. Sin embargo, su lectura requiere una preparación suficiente, puesto que, como señalan los autores en el prefacio, para su comprensión es necesario un conocimiento preciso de los fundamentos de la mecánica cuántica, de su aplicación al enlace químico, de la estadística, tanto clásica como cuántica, y del modelo de gas de Fermi. En este sentido el libro podría ser asequible a un mayor sector de estudiosos si se hubiera incluido en el mismo una introducción suficientemente amplia acerca de los temas señalados.

La presentación es muy agradable. Como comentario final señalaremos que su aparición resulta muy oportuna, por las siguientes razones; *a*) los recientes avances en la superconductividad a temperaturas elevadas (párrafo 8.8); *b*) la controvertida fusión en frío en el seno del Pd (párrafos 3.7 y 5.8.1) y *c*) la recentísima obtención del hidrógeno sólido, que sería, estructuralmente hablando, el metal más simple y, por consiguiente, el de más fácil tratamiento teórico. (L. G.)

**TÉCNICA Y PODER EN CASTILLA DURANTE LOS SIGLOS XVI Y XVII**, por Nicolás García Tapia. Junta de Castilla y León, 1989.

La zona norte de Castilla, que comprende aproximadamente la cuenca española del Duero, fue el centro de la política, la economía, el comercio y, desde luego, la técnica. Hasta el asentamiento definitivo de la corte en Madrid, la zona norte de Castilla fue el foco de una actividad de ingeniería que se diferenció netamente de la del resto de la península. No quiere esto decir que no hubiera influencias, singularmente con la interesante ingeniería aragonesa, sino que la técnica castellana cobró una personalidad propia que se reflejó en importantes realizaciones.

Si ahora nos parece evidente la influencia de la técnica en la historia, ¿por qué negarla para una época caracterizada por profundos cambios en la tecnología? Muchos libros de historia pasan por alto los cambios técnicos que tuvieron lugar en el siglo XVI, como si no hubieran sido fundamentales en los acontecimientos históricos. Entonces, como ahora, poder y técnica iban unidos, y una falta de atención hacia los avances de la tecnología significaba la decadencia del poder.

Los encargados de la invención y producción técnica eran los ingenieros, que entonces empezaban a formarse como grupo profesional. Los ingenieros, surgidos por escisión de los arquitectos, eran los autores también de las obras públicas: puentes, caminos, acueductos, abastecimientos de agua, etcétera, que formaban el entramado imprescindible para la buena marcha del conjunto del territorio. Los conocimientos técnicos eran necesarios para la realización de estas obras que eran patrocinadas por la corona, los municipios o los señores eclesiásticos o nobles.

En Castilla, en los siglos XVI y XVII ya existían las vías de intercambios y producción de tecnología que caracterizan nuestro actual sistema industrial. Estas son la importación de tecnología, el llamado "espionaje industrial" y la invención propia. De estos tres procedimientos de adquisición de ingeniería, este libro presenta varios ejemplos notables que se dieron entonces en la región castellana. Desde finales del siglo XV, y como consecuencia de la estancia de la corte en diversos lugares de la meseta, ésta se convirtió en la espina dorsal sobre la que se sustentaron los esfuerzos de la renovación de todo el territorio español.

El ingeniero polifacético, artista, científico, técnico y humanista, surgió como una necesidad del Renacimiento, primero en Italia (Leonardo da Vinci) y luego en las cortes europeas. Si bien se ha estudiado el caso de estos ingenieros italianos renacentistas, no es bien conocida la existencia de un nutrido plantel de ingenieros españoles que reunieron las características de los italianos e incluso les superaron en muchos aspectos. Los nombres de Jerónimo Givarra, Pedro Juan de Lastanosa, Pedro de Esquivel, Benito de Morales y otros son prácticamente desconocidos, aunque superaron como ingenieros a los mejores de esta profesión en su época. Determinados arquitectos como Juan Bautista de Toledo y Juan de Herrara han sido muy bien estudiados como excelentes arquitectos, pero su labor como ingenieros, igualmente notable, es punto menos que ignorada.

Los principales temas abordados por este libro se pueden resumir en las vías de comunicación, la ingeniería hidráulica, la industria artesanal, las fuentes de energía... Referente a las comunicaciones cabe recordar que durante la Edad Media se habían abandonado gran parte de las calzadas y puentes construidos en época romana, pero las necesidades del transporte militar en la Reconquista y la incidencia de las vías de peregrinación, singularmente el Camino de Santiago, aconsejaron una política de reparación y construcción de una red de infraestructura viaria. Había que salvar los ríos por puentes, al principio sobre barcas de madera y luego construidos de forma definitiva en sólida cantería para evitar los gastos frecuentes de reparaciones por las crecidas. Los escritos renacentistas que se han conservado sobre puentes nos enseñan sobre la forma en que se construían. Para el siglo XVI, el mejor texto que nos ilustra sobre los puentes es el de *Los veintiún libros de los ingenieros y de las máquinas*.

En tiempo de los Reyes Católicos la ingeniería hidráulica vivió un inusitado resurgimiento, pues se incrementó la construcción de acequias y canales en España. Arquitectos e ingenieros como Juan Bautista de Toledo, Jerónimo Givarra, Pedro Juan de Lastanosa, Pedro de Esquivel, Juan de Herrera y otros muchos, entre ellos italianos como Juan Francisco Sitoni y un insólito personaje, el ermitaño-ingeniero Mariano Azaro, fueron contratados por Felipe II para intervenir en obras de acequias y canales de regadío.

En Castilla y León, ya desde el siglo XV se empezó a utilizar la técnica del abastecimiento de agua de poblaciones, palacios y monasterios, con arreglo a una nueva mentalidad de disfrute y de higiene. Por ejemplo, el Monasterio de Santa María del Parral en Segovia poseía, desde mediados del siglo XV, un abastecimiento propio para sus huertas y jardines hasta el punto de haberse acuñado el dicho de que "los huertos del Parral son el paraíso terrenal". Esta afirmación indica la idea extendida, a finales de la Edad Media y principios de Renacimiento, de tratar de reproducir, aquí, el paraíso terrenal perdido.

Para conservar el agua en un edificio, se recurría a depósitos que recibían el nombre de *cisternas* cuando recogían el agua de lluvia y *aljibes* si eran depósitos de acumulación de agua procedente de una conducción. Cuando una cisterna recogía el agua de lluvia que caía sobre los tejados, se acumulaba en un depósito de forma cilíndrica que tenía dobles paredes; primero pasaba el agua a la cámara interior y se filtraba, a través de la pared cilíndrica, hasta llegar depurada al depósito exterior, desde donde salía través de un orificio practicado en su parte más baja. Si el agua que caía sobre los tejados no era suficiente para la alimentación de la casa, se recurría a un sistema formado por un depósito al aire libre que recogía el agua de lluvia y de los barrancos. El agua se filtraba a través de una capa de arena que estaba sobre la bóveda de un depósito subterráneo; éste tenía una boca, como el brocal de un pozo, por donde se extraía el agua con cubos o norias de cangilones. Este tipo de cisterna era muy utilizado en edificaciones rurales y palacios, ya que el agua se conservaba fresca y agradable, añadiéndole ciertas sustancias para su depuración.

La industria artesanal castellanoleonesa se centra, en este libro, en la industria del hierro. Con la llegada de las invasiones indoeuropeas, que penetraron por las estribaciones occidentales de los Pirineos, comienza la fabricación y utilización del hierro, durante la llamada Primera Edad del Hierro. Se partía del mineral del óxido férrico, que se reducía, por medio del carbón vegetal, en pequeños hornos de dos o tres metros de altura y uno de diámetro. Mediante el martillado de la masa de hierro obtenida, se eliminaban las escorias, dando forma a los objetos de hierro. El sistema permaneció básicamente sin cambios hasta la llegada de los romanos, quienes introdujeron una serie de mejoras, tanto en los hornos como en la producción de objetos de hierro. En primer lugar cambió el tipo de horno rudimentario de los celtas por uno más perfeccionado, que recuerda los actuales de *reverbero*. En segundo lugar, comienza una especialización de funciones en la producción de hierro, que parte de la *minería* o técnica de extracción de mineral y llega a la *fundición* y a la *forja*.

Las técnicas de producción de hierro seguirán casi inalteradas en los primeros tiempos de la reconquista cristiana hasta la introducción de la *farga catalana*, que superó los hornos de tipo reverbero de época romana y visigótica. El sistema procedía del norte de la Península y se extendió a todo el territorio hispano e incluso a Europa. El procedimiento de las fraguas catalanas tuvo posiblemente su antecedente en las primitivas ferrerías inglesas y prusianas. Fueron perfeccionadas en tierras catalanas, de ahí el nombre con que se conocen en todo el mundo.

La ferrería más conocida de la zona castellano-leonesa es la de Compludo, de la que está documentada su existencia en el siglo VII. Sin embargo, la maquinaria que todavía se contempla puede proceder de los siglos XVI y XVII, aunque con algunas modificaciones. La energía para el movimiento de todo el conjunto la proporciona el agua de los arroyos de Miera y Miruelos, que se canaliza por una conducción abierta llamada *antepara*, la cual lleva a un depósito de acumuación de agua que en Compludo está excavado en la roca y se denomina *banzao*. La misión de este depósito es que el agua obtenga cierto nivel para que adquiera la suficiente energía potencial como para hacer girar la rueda hidráulica; al mismo tiempo, el depósito regula el caudal de aporte y acumula el agua para cuando tenga que funcionar. En este sentido, su misión es la misma que la del *cubo* de los molinos hidráulicos.

La rueda hidráulica es vertical (de eje horizontal), fabricada enteramente en madera. Sobre la periferia de la rueda, se encajan unos rudimentarios álabes rectos, también en madera. Los radios de la rueda son solidarios de un grueso eje formado por un tronco de unos 70 cm de diámetro. La rueda hidráulica está fuera del edificio, pero se comunica por medio de un estrecho arco, que permite acceder directamente al exterior, sin rodear la fábrica. Por este acceso pasa el eje, que termina en dos piezas cilíndricas que le sirven de cojinete de giro. El eje va zunchado por aros de hierro en sus puntos débiles.

Uno de los factores más indicativos de la marcha de una sociedad industrial es el uso de las fuentes de energía disponibles. Durante el Renacimiento se dejó sentir la necesidad de disponer de una fuente de energía cómoda, barata y transportable. La madera y el carbón sólo eran utilizados como combustible doméstico o para el funcionamiento de hornos en aplicaciones metalúrgicas, no como fuente energética para mover máquinas, a falta de la invención del motor apropiado. Es importante saber cómo se utilizaban las fuentes de energía en forma de cursos de agua o de corrientes de viento, para el funcionamiento de las incipientes industrias renacentistas. En épocas anteriores a la Revolución Industrial, el papel del "molino" era equivalente al de nuestra actual factoría.

Los molinos hidráulicos constituyeron durante siglos un elemento esencial en el paisaje de los ríos castellanos. Las grandes ruedas girando con la fuerza del río o del canal o *caz* fueron el símbolo de la potencia

adquirida por las industrias que sobrepasaban la condición de artesanales. La Edad Media supuso el desarrollo y la generalización del molino hidráulico, en parte por tradición árabe y en parte también como producto del ingenio de los constructores españoles. Su técnica avanzada hace que tengamos que diferenciar varios tipos de molinos de agua, en función del modo de aprovechamiento de la energía hidráulica, cuyo elemento esencial es la *rueda hidráulica*, que puede ser de varios tipos.

El libro que comentamos termina descubriendo la unión de la técnica alemana con la arquitectura española. La Casa de la Moneda de Segovia, emplazada a orillas del Eresma al lado del puente del Parral, es uno de los edificios industriales más importantes de la época renacentista en la región de Castilla y León. En la obra arquitectónica intervinieron Juan de Herrera y su sucesor Francisco de Mora, junto a los mejores aparejadores de su tiempo.

Las máquinas de acuñación de moneda de Segovia constituyen los primeros ingenios de este tipo de España que utilizaron la energía hidráulica aplicada al proceso de laminación de la moneda, lo que supuso una importante revolución tecnológica al conseguir una rápida acuñación de la misma, necesaria en el momento económico que vivía la España de Felipe II.

Sin embargo, el proceso técnico utilizado en la Casa de la Moneda de Segovia no fue una invención española y constituye un ejemplo de importación de tecnología, al emplear a especialistas alemanes para el montaje de una máquina cuya construcción fue realizada por entero en Alemania, aunque, en realidad, el invento de las máquinas hidráulicas para acuñar moneda no es alemán.

Este es, a grandes rasgos, la temática de este libro tan interesante cuan ameno para los aficionados a la historia. Un libro que, sin reservas, recomendamos. (P. M.)

**BUFFON. LES ÉPOQUES DE LA NATURE**. Edición crítica de Jacques Roger. Editions du Muséum; París, 1988. **MAN, GOD, AND NATURE IN THE ENGLIGHTENMENT**. Dirigido por Donald C. Mell, Jr., Theodore E. D. Braun y Lucia M. Palmar. Colleagues Press; East Lansing, 1988.

Si de alguien puede decirse, abrazando la paradoja, que influyó en todos y en nadie, ese fue Georges-Louis Leclerc, conde Buffon (1707-1778). Su verbo exacto y sugestivo, que se apegaba a la observación directa y dejaba libre la fantasía, y el nervio de su estilo, que ganaba al lector, le atrajeron un sinnúmero de devotos, aunque ninguno le siguiera. Quizá por la razón principal de que él mismo apenas tuvo una idea asentada y cambió de pensamiento con la fugacidad de las estaciones. Se opuso, con distante altivez, lo mismo a Linneo que a los antagonistas de la geología neptuniana. Sus aficiones sistemáticas –globalizadoras– eran propias de otro tiempo, mientras que traía nuevos aires provocadores en sus expresiones de doble sentido, de regusto por amagar y no dar; a causa de lo primero, pocos se atrevían a citarle; no obstante lo segundo, su hijo murió en la guillotina.

Del pensamiento de Buffon se ha ocupado, en repetidas ocasiones y variada profundidad, Jacques Roger desde sus primeros trabajos historiográficos. Este de la edición crítica de *Les Époques de la nature* es un estudio de mocedad, editado en 1962 y reimpreso en 1988 para conmemorar el segundo centenario de la muerte del naturalista. Con una óptica algo diferente, mayor hondura y menos apasionamiento, John Lyon y Phillip R. Sloan publicaron en 1981 el que, a mi juicio, constituye la investigación más seria acometida hasta hoy: *From Natural History to the History of Nature: Readings from Buffon and His Critics*. Agrega ésta al final la traducción al inglés del *Voyage à Montbard* de Hérault de Séchelles, un cuadro más o menos periodístico de la figura de Buffon que ha servido de punto de referencia para todas las aproximaciones al Intendente del Jardín Real de París. (Ese relato frívolo con pretensiones que la historia ha puesto en solfa fue traducido en parte al español en 1797; pero los fragmentos eliminados son tantos, y de capital interés, que arruinaron el esfuerzo.)

Aunque *Les Époques de la nature* lleva pie de imprenta de 1778 no salió al comercio hasta la primavera del año siguiente. Viene a ser como el testamento de la copiosa obra de Buffon; recuérdesele autor de una famosa *Histoire naturelle* en 36 volúmenes, un suplemento a la misma, que recoge también, reescritas, sus comunicaciones a la Real Academia de Ciencias, y otros libros de valor dispar. Su campo de interés es universal y sus escrúpulos, a la hora de apropiarse de manuscritos ajenos, nulos.

Roger recrea la aventura científica de Buffon en torno a la concepción, gestación y redacción (varias) de las *Époques*. Tratando estas principalmente de la formación planetaria y de la aparición y distribución de las especies, los puntos de ligazón con la precoz *Théorie de la terre* (1749) son múltiples y, valiente, su adscripción newtoniana. La explicación de conchas fósiles halladas en regiones elevadas constituyó, sabido es, un motivo desencadenante de la investigación paleontológica y de las teorías sobre la historia remota de la Tierra. Buffon, desde su puesto de Intendente del Jardín Botánico de París, se encontraba en la mejor situación para conocer lo observado por distintos expedicionarios (La Condamine, por ejemplo), pero le faltó humildad para supeditar la teoría neptuniana a los datos que negaban la arribada de las aguas a los picos andinos. Ingeniosas fueron, empero, sus hipótesis sobre la distribución de las especis en América y sobre la relación entre cambios climáticos y tamaño de los individuos.

Que la corriente fluía en otro sentido es cosa aceptada en la historiografía del XVIII, aunque el curso de las ideas no sea nunca unilateral y coexistan tendencias contrapuestas o mixtas en cada momento y, a veces, en un mismo autor. El propio Roger es uno de los que ha insistido con mayor énfasis en el vuelco de los nuevos planteamientos: se abandonan los grandes sistemas, las teorías generales, llámese cartesianismo, llámese anacrónicamente vitalismo, para insistir en los hechos observados. De ese nuevo rumbo da cumplida nota *Man, God, and Nature in the Enlightenment*, compendio de artículos sobre distintos aspectos del siglo de las luces: filosofía, naturaleza, sociedad, religión (este apartado cuarto cubre, en realidad, religión, moral y literatura) y ciencia.

La literatura reclama alguna que otra vez la atención del historiador de la ciencia (Picatoste publicó un magnífico estudio, desde esa óptica, sobre Calderón de la Barca) y, más frecuentemente, la filosofía y la religión. Aquí, sin embargo, nos limitaremos al quinto apartado: la ciencia propiamente dicha en el siglo XVIII. Abarca cuatro artículos sin relación buscada. Virginia P. Dawson estudia el epistolario de Charles Bonnet sobre la partenogénesis en los áfidos. Su descubrimiento de la reproducción "sans accouplement" desmentía el principio incontrovertido según el cual se necesitaba, en toda generación, el concurso de los dos sexos. Con ser importante el resultado, yo destacaría la ambientación que se reconstruye; es decir, la importancia de la comunicación epistolar entre los científicos que permitía detenerse en los fracasos, el método, en las ideas que guiaban la experimentación. Las cartas de ese tenor escritas por Bonnet ocupaban cuantiosas páginas, en letra piojosa, y, aunque se dirigían a personas

determinadas (Abraham Trembley o René-Antoine Réamur), se esperaba que se presentaran ante sociedades o academias que dieran respaldo institucional al emisor.

En repetidas ocaciones se ha abordado un fenómeno curioso, hoy no suficientemente explorado, o quizá reputado artificial; ¿qué relación guardan el origen de los lenguajes y la taxonomía o clasificación biológica? Una obra imprescindible es la de M. M. Slaughter, *Universal languages and scientific taxonomy in the seventeenth century* (1982). W. Keith Percival aplica ese enfoque al siglo siguiente, si bien no aporta nada nuevo, salvo la buena selección de las notas. En un orden complementario, Bonnie Arden Robb explica la superioridad, *Diderot teste*, del francés sobre el inglés para designar la naturaleza de los objetos a través de la sucesión ordenada de los adjetivos. Joseph Musser, Jr., por último escribe un artículo claro de metaciencia donde saca a la luz las limitaciones del modelo paradigmático de Thomas Kuhn; de paso aduce ejemplos ilustradores de cómo unos mismos experimentos posibilitaron diferentes interpretaciones. O lo que es lo mismo, que el dato desnudo no sirve para nada, si no va acompañado de una teoría o una hipótesis. Una suerte de victoria pírrica de Buffon. (L. A.)

**QUANTUM THEORY AND PICTURES OF REALITY**, dirigido por W. Schommers; Springer-Verlag; Berlín-Heidelberg, 1989.

En 1935 Einstein, que acababa de incorporarse al Instituto de Estudios Avanzados de Princeton, publicó, junto con su antiguo colaborador Boris Podolsky y el joven, recién doctorado, Nathan Rosen, un artículo con un desafiante título: "¿Puede la descripción de la mecánica cuántica de la realidad física considerarse completa?" (*Physical Review* 47, 777 (1935)). Si bien el debate sobre los fundamentos de la teoría cuántica había comenzado antes (en 1927, en Como, con las discusiones entre Bohr y Einstein), puede afirmarse que el citado artículo constituye la primera publicación en donde se exponen con rigor y precisión los aspectos problemáticos de la mecánica cuántica.

Así lo advirtió Bohr, quien no tardó en responder a los argumentos presentados en ese artículo, aunque con no excesiva convicción. De todos modos, la discusión se planteó de modo más "físico", más conectada con el mundo experimental, a partir de los importantes trabajos de Bell en 1964-65. Las desigualdades deducidas por Bell, que distinguen la mecánica cuántica de cualquier teoría con elementos objetivos de realidad local (del tipo propugnado por Einstein), han sido estudiadas en el laboratorio repetidas veces, con resultados claramente a favor de las predicciones de la mecánica cuántica.

Por su reacción a los argumentos de Einstein y, en general, a la cuestión de la realidad –y la separabilidad– en la teoría cuántica, los físicos pueden clasificarse, según Mermin, en dos tipos: *1*, los preocupados por el tema; *2*, la mayoría, que no lo están. Pero el tipo *2* puede aún dividirse en dos subtipos: *2a*, los físicos que explican por qué no están preocupados. Sus explicaciones tienden bien a malinterpretar frecuentemente los puntos relevantes en el argumento EPR, o bien a contener asertos de tipo físico que pueden demostrarse falsos. En cuanto a los encuadrados en el subtipo *2b*, ni se sienten preocupados ni dan explicaciones del por qué de su desinterés. (Quizá fuera de ambos tipos debamos colocar a Feynman, que decía al respecto: "No me siento capaz de definir el problema real, así que sospecho que no hay ningún problema real, pero no estoy seguro que no haya ningún problema real".)

El editor de este libro y los que en él contribuyen han de clasificarse en el tipo *1*. Lo que se discute es justamente el concepto de realidad física y las imágenes de ésta a la luz de la teoría cuántica contemporánea.

El libro es una recopilación de artículos de diferentes autores. El editor contribuye con dos: el primero, "Evolution of Quantum Theory" (que es también el primero del libro) es, como indica su título, un estudio de tipo histórico-crítico sobre el desarrollo de la teoría cuántica. Aunque está bien escrito, este artículo tiene sólo interés "pedagógico" (que, por otra parte, no es poco); mucho de original no puede encontrarse en él. En su segundo artículo, "Space-Time and Quantum Phenomena", Schommers analiza con pormenor el problema del espacio y, sobre todo, del tiempo en la mecánica cuántica. Me parece ante todo digno de resaltar su amplio resumen de la teoría de Prigogine y colaboradores, que, aunque todavía incompletamente elaborada y discutible en algunos de sus puntos, puede sin duda verse como una interesante aproximación al problema de la complejidad estructural y su relación con el indeterminismo cuántico.

El segundo de los ensayos lo firma P. Eberhard y lleva por título "The EPR paradox, Roots and Ramifications". A partir de las correlaciones EPR, se analiza la eventual propagación de señales superlumínicas en virtud de tales correlaciones instantáneas. Aparte del punto de vista ortodoxo (o convencional), en el que dicha posibilidad no tiene cabida, se describen algunos otros modelos –como el de Costa de Beauregard de propagación hacia atrás con el tiempo, más propio de la ciencia-ficción que de otra cosa– en los que tal tipo de señales serían en principio posibles. Eberhard expone su propio modelo en otro de los capítulos del libro, el cuarto, desde mi punto de vista menos interesante.

El ensayo de más entidad es el que constituye su capítulo tercero, cuyo autor es B. d'Espagnat, y está basado en un artículo de éste aparecido en *Physics Reports* en 1984. Su título, "Nonseparability and the Tentative Descriptions of Reality", da ya una idea de por dónde van las cosas. Está escrito con autoridad y rigor, y en él se examinan las implicaciones sobre todo, pero no sólo, filosóficas de los aspectos de inseparabilidad inherentes a la teoría cuántica.

El libro se cierra con una contribución de F. Selleri, "The Wave-Particle Duality: Recent Proposals For The Detection of Empty Waves", en la que se aborda, apoyándola, la idea debida a Einstein y de Broglie de la existencia de ondas vacías, o camposfantasma ("Gespenstenfelder") según la terminología de Einstein, dotados de realidad objetiva pero que no propagan, de acuerdo con su nombre, ningún tipo de energía. Es decir, es un intento, infructuoso, de dotar de alguna manera de realidad a la teoría cuántica sin cambiar (y más vale que así ocurra) sus predicciones estadísticas, al menos en lo que concierne a los experimentos realizados hasta la fecha, que han sido satisfactoriamente explicados por dicha teoría. El lector interesado en este tipo de ideas puede consultar del mismo Selleri "*El debate de la teoría cuántica*" en cuya parte final se presenta una exposición de lo esencial de ellas. En realidad, la contribución al libro que aquí se comenta es una versión, bastante ampliada y con mayor énfasis en los aspectos de tipo experimental.

En resumen, a pesar de que la bibliografía sobre el problema de la realidad en teoría cuántica a la luz del aparente subjetivismo y la indiscutible inseparabilidad (y, por tanto, en algún sentido, no-localidad) de ésta, empieza a ser algo excesiva, no hay duda de que todo libro que aporte algo original, o sirva para entender mejor lo esencial del problema aunque no sea de gran originalidad, debería ser bien recibido. Creo que este libro pertenece a esta última categoría. (J. L. S.)

# Apuntes

Los virus son parásitos intracelulares. El agente antivírico ideal sería, pues, aquel que impidiera su entrada en la célula huésped, lo que mermaría grandemente el daño infligido por la agresión y por la respuesta inmune desencadenada contra la misma. La inmensa mayoría de los agentes antivíricos disponibles ($\alpha$-interferón, aciclovir, AZT y ribavirin) actúan sólo una vez invadida la célula, franqueado el paso por sus receptores de superficie a los que se anclan los agentes. Se conoce ya, un primer paso necesario, la identidad molecular de algunos de esos caballos de Troya. La esperanza se ha afianzado con el descubrimiento de un análogo de los rhinovirus que se engarza con los receptores de los virus del resfriado común.

En España el grueso de la investigación y de la actividad universitaria pertenece al Estado. Una situación cuya tendencia se está acentuando con las medidas legales tomadas en los últimos años. No parece ser ese el camino del desarrollo, si nos espejamos en Japón, por ejemplo, donde el 80 por ciento de la investigación se realiza en la empresa privada. Más aún. Permeables a las nuevas ideas y bagaje que puedan traer expertos del extranjero, las universidades privadas niponas contrataron, sólo en 1987, 392 catedráticos occidentales, frente a la docena escasa que entró en la pública.

Entre bromas y veras se dice del entomólogo que es una especie a extinguir. Por lo menos está poco protegida, si echamos una ojeada a las ayudas a su investigación, becas, plazas y demás subvenciones. Lo paradójico del caso es que ese declive coincide con la creciente importancia que los insectos están adquiriendo en la vida moderna. Por culpa de esos invertebrados, se perdía, en 1940, el 7 por ciento de las cosechas; en 1980 arrasaron con el 14 por ciento. El DDT no acabó con la malaria, pero está matando los malariólogos, sentencian en el sector.

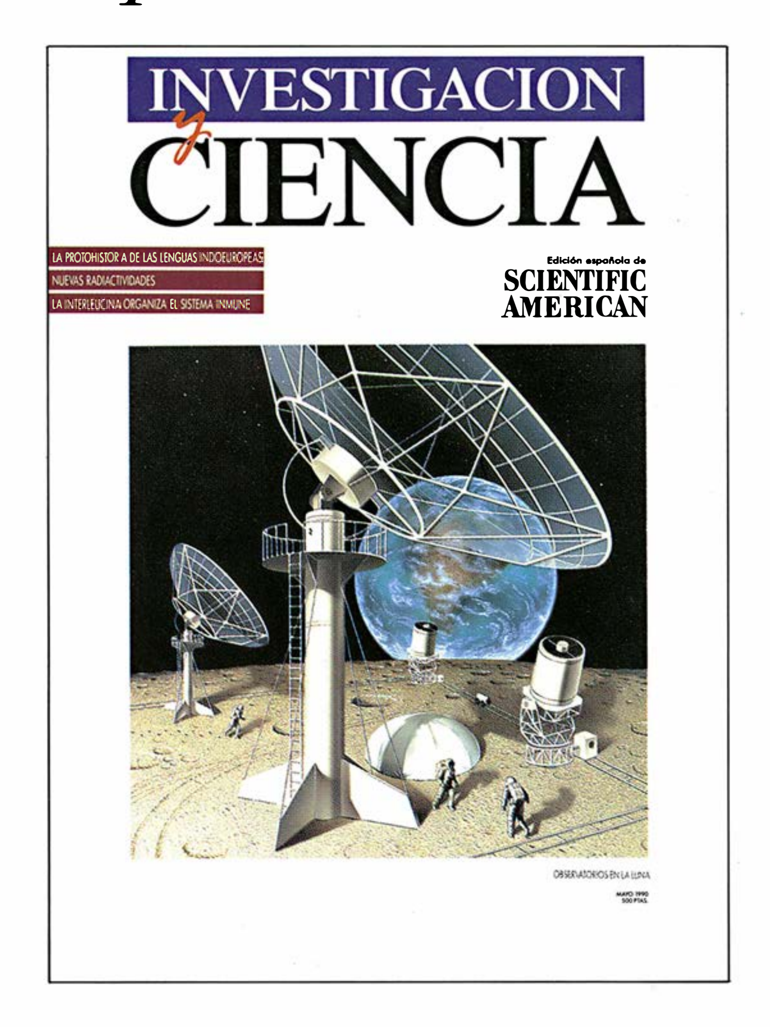
Se creía que los briófitos constituían los únicos vegetales terrestres que desconocían la alopoliploidía, proceso de extraordinario interés evolutivo y en el que se hallan implicadas la hibridación entre especies distintas y la duplicación cromosómica. Los datos obtenidos por electroforesis acaban de descubrir un contraejemplo que cuestiona la tesis general: el musgo poliploide *Plagiomnium medium* es un derivado alopoliploide de *P. ellipticum* y *P. insigne*. Eso no es todo. *P. medium* se ha originado más de una vez de estos progenitores, fecundación cruzada que resulta en recombinación genética.

Las células normales poseen, en un medio de cultivo, limitada su capacidad de proliferación. En el caso de los fibroblastos, su esperanza de vida varía de 20 a 60 duplicaciones. Los fibroblastos senescentes no mueren en una fase quiescente, sino que presentan alterada su expresión génica. De siete genes estudiados, cuatro se expresaron con más o menos normalidad en las células seniles, si bien uno de ellos no produjo la cuantía necesaria de la enzima que codificaba; el rendimiento de los otros tres genes quedó notablemente mermado, así el del oncogén *c-fos*, cuya expresión normal se considera esencial para la división celular.

Para limpiar la contaminación producida por la práctica ritual de los hindúes de arrojar los cuerpos semiquemados al Ganges, las autoridades han pensado en repoblar el río con tortugas carnívoras. En una estación reptiliana están ya preparados 7000 ejemplares que llevarán a cabo la labor que no pueden acometer los 32 crematorios eléctricos dispuestos a lo largo del curso, rechazados por la conciencia de los hindúes practicantes.

Algunas teorías sobre partículas elementales predicen la existencia de líneas finísimas de energía concentrada, las cuerdas cósmicas. Se formarían en una transición de fase inmediatamente después de la gran explosión inicial y las encontraríamos quizás hoy en regiones remotas del universo. No se ha descubierto todavía la radiación gravitatoria que probara la existencia de tales cuerdas. Antes bien, y en oposición a la evolución de escala invariante que se les suponía, las últimas simulaciones apuestan por variaciones de las cuerdas en forma de ángulos agudos, que les daría una configuración zigzagueante en los puntos de trabazón entre cuerdas. Esos dobleces se irían distendiendo con la expansión del universo, morfología que, al propio tiempo, propiciaría una medición más fácil de la radiación gravitatoria y, por tanto, la confirmación de la existencia de las cuerdas.

# *Seguiremos explorando los campos del conocimiento*




**BIODEGRADACION E INGENIERIA GENETICA, por Juan Luis Ramos y Fernando Rojo**
*Las bacterias pueden degradar una amplia variedad de compuestos orgánicos. La manipulación genética de sus rutas catabólicas es una vía eficaz para acelerar su evolución y permitir la mineralización de compuestos.*

**INTERLEUCINA-2, por Kendall A. Smith**
*Primera hormona del sistema inmune en ser reconocida, auxilia al cuerpo a preparar la estrategia de defensa contra la invasión de microorganismos.*

**TERREMOTOS EN LA CORTEZA CONTINENTAL ESTABLE, por Arch C. Johnston y Lisa R. Kanter**
*Los terremotos pueden producirse incluso en una corteza estable, lejos de las zonas sísmicas situadas en los bordes de las placas tectónicas. ¿Cómo explicar tan enigmáticos sucesos?*

**NUEVAS TENDENCIAS EN LA COMUNICACION, por Karen Wright**
*¿Qué piensan la empresa y los científicos sobre las innovaciones que están cambiando nuestra vida moderna con la introducción de los medios informáticos en el quehacer y en el ocio diarios?*

**VERTEBRADOS SUSPENSIVOROS, por S. Laurie Sanderson y Richard Wassersug**
*Los animales que obtienen su alimento filtrándolo del agua pueden recolectar la abundancia del plancton y crecer en gran número o alcanzar un tamaño imponente.*

**NUEVAS RADIACTIVIDADES, por Walter Greiner y Aurel Sandulescu**
*Un núcleo puede restructurarse espontáneamente desprendido ocasionalmente agregados extraños de protones y neutrones. Las observaciones de estas nuevas radiactividades han arrojado nueva luz sobre las teorías de la dinámica nuclear.*

**EL FUNCIONAMIENTO DEL INCONSCIENTE, por Joseph Weiss**
*Rigurosos estudios cuantitativos de psicoterapia están poniendo en tela de juicio algunas opiniones muy generalizadas acerca de cómo trabaja el inconsciente y cómo progresan los pacientes sometidos a terapia.*

**LA PROTOHISTORIA DE LAS LENGUAS INDOEUROPEAS, por Thomas V. Gamkrelidze y V. V. Ivanov**
*Según dice el autor, los rastros del antepasado común de estas lenguas han sido hallados en Asia y no en Europa. La otrora clara distinción entre las ramas oriental y occidental de esta familia se ha vuelto hoy borrosa.*

